



N/Alternate names: cysteine proteinase inhibitor; onchocerciasis antigen  
 C/Species: Onchocerca volvulus  
 C/Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 17-Mar-2000  
 C/Accession: A43428; B43927  
 R/Author: S.; Brozman, B.; Huima, T.; Prince, A.M.; Mckerrow, J.H.  
 J. Biol. Chem. 267, 17339-17346, 1992  
 A/Title: Molecular cloning and characterization of onchocystatin, a cysteine proteinase  
 A/Reference number: A43428; MUID:92381053; PMID:1512269  
 A/Accession: A43428  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-162 <IUS>  
 A/Cross-references: GB:M37105; NID:g159905; PID:g159906  
 A/Note: sequence extracted from NCBI backbone (NCBIN:111962, NCBIPI:111963)  
 R/Author: R.; Maasod, K.; Alvarez, R.M.; Ogunrinde, A.F.; Lujan, R.; Richards J.  
 J. Clin. Invest. 88, 1460-1466, 1991  
 A/Title: Molecular cloning and characterization of recombinant parasite antigens for imm  
 A/Reference number: A43927; MUID:92042723; PMID:1840605  
 A/Accession: B43927  
 A/Molecule type: mRNA  
 A/Residues: 'P', 37-57, 'A', 59-71, 'R', 73-83, 'N', 85-126, 'W', 128-162 <CHA>  
 A/Cross-references: GB:M60279; NID:g159888  
 A/Experimental source: clone OC 9.3  
 A/Note: sequence extracted from NCBI backbone (NCBIN:65111, NCBIPI:65113)  
 C/Superfamily: cystatin from NCBI homology  
 C/Keywords: cysteine proteinase inhibitor

Query Match 39.2%; Score 56; DB 2; Length 162;  
 Best Local Similarity 45.5%; Pred. No. 0.51;  
 Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 4 NKESDDKHFRRFLVKYQROV 25  
 DB 77 NEQSDNEYHMPKILKVSQV 98

# RESULT 12

UDHUP1

cystatin S precursor - human

N/Alternate names: cystatin SA-II; salivary acidic protein-1

C/Species: Homo sapiens (man)

C/Date: 25-Feb-1985 #sequence\_revision 08-Feb-1996 #text\_change 16-Jul-1999

C/Accession: S17667; S16500; A01272; A29603; S19280; A56608

R/Author: L.A.; Aguirre, A.; Levine, M.J.

Biochem. J. 278, 627-635, 1991

A/Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ an

A/Reference number: S17667; MUID:91378918; PMID:1898352

A/Accession: S17667

A/Molecule type: mRNA

A/Residues: 1-141 <BOB>

A/Cross-references: EMBL:X54667; NID:g30365; PIDN:CAA38478.1; PID:g30366

R/Author: M.S.; Jensen, J.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.

Arch. Biochem. Biophys. 288, 664-670, 1991

A/Title: Salivary cystatin SA-II, a potential precursor of the acquired enamel pellicle

A/Reference number: S16500; MUID:91378515; PMID:1898055

A/Accession: S16500

A/Status: preliminary

A/Molecule type: protein

A/Residues: 21-134, 'D', 136-141 <IND>

R/Author: S.; Saitoh, E.; Sanada, K.

J. Biochem. 96, 489-498, 1984

A/Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole sal

A/Reference number: A91985; MUID:85054716; PMID:6501254

A/Accession: A01272

A/Molecule type: protein

A/Residues: 29-134, 'D', 136-141 <ISE>

R/Author: S.; Saitoh, E.; Ito, S.; Isemura, M.; Sanada, K.

J. Biochem. 96, 1311-1314, 1984

A/Title: Cystatin S: a cysteine proteinase inhibitor of human saliva.

A/Reference number: A91981; MUID:85104877; PMID:6334600

A/Contents: annotation; inhibitor specificity

R/Author: D.H.; Yuan, P.M.; Wilson, K.J.; Hunkapiller, M.W.

Biochem. Biophys. Res. Commun. 145, 1248-1253, 1987

A/Title: Identification of a long form of cystatin from human saliva by rapid microbo  
 A/Reference number: A29603; MUID:87270697; PMID:3496880  
 A/Accession: A29603  
 A/Molecule type: protein  
 A/Residues: 21-51 <HAW>  
 R/Author: N.; Reddy, M.S.; Bergey, E.J.; Haraazthy, G.G.; Soni, S.D.; Levine, M.U.  
 Biochem. J. 280, 341-352, 1991  
 A/Title: Large-scale purification and characterization of the major phosphoproteins and  
 A/Reference number: S19279; MUID:92082469; PMID:1747107  
 A/Accession: S19280  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 21-55 <RAM>  
 R/Author: M.; Richardson, C.F.; Bergey, E.J.; Levine, M.J.; Nancollas, G.H.  
 Arch. Oral Biol. 36, 631-636, 1991  
 A/Title: The effects of human salivary cystatin and statherin on hydroxyapatite crystal  
 A/Reference number: A56608; MUID:92074899; PMID:1741693  
 A/Accession: A56608  
 A/Molecule type: protein  
 A/Residues: 21-36 <JOH>  
 A/Note: sequence extracted from NCBI backbone (NCBIPI:67866)  
 C/Comment: This protein strongly inhibits papain and ficin, partially inhibits stem brom  
 competitively.

A/Genes: GDB:CST4

A/Cross-references: GDB:136381

A/Map position: 20p11.2-20p11.2

C/Superfamily: cystatin; cystatin homology

C/Keywords: cysteine proteinase inhibitor; phosphoprotein; saliva

F/1-20/Domain: signal sequence #status predicted <SIG>

F/21-141/Product: cystatin S #status predicted <MAT>

F/30-141/Domain: cystatin homology <CYS>

F/76-80/Region: inhibitory #status predicted

F/94-104,118-138/Disulfide bonds: #status predicted

Query Match 37.8%; Score 54; DB 1; Length 141;  
 Best Local Similarity 39.1%; Pred. No. 0.88;  
 Matches 9; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 2 QYNKESDDKHFRRFLVKYQRO 24  
 DB 54 EYNKATEDERYRRLQVLRAREQ 76

# RESULT 13

UDHUP2

cystatin SN precursor [validated] - human

N/Alternate names: cystatin SA-I

C/Species: Homo sapiens (man)

C/Date: 28-May-1986 #sequence\_revision 08-Feb-1996 #text\_change 08-Dec-2000

C/Accession: A28110; S02489; A29632; A01273; S19279

R/Author: I.; Dickinson, D.P.; Levine, M.J.

J. Biol. Chem. 263, 9381-9387, 1988

A/Title: Purification, molecular cloning, and sequencing of salivary cystatin SA-I.

A/Reference number: A28110; MUID:88243825; PMID:2837486

A/Accession: A28110

A/Molecule type: mRNA

A/Residues: 1-141 <ALH>

A/Cross-references: GB:003870; NID:g337751; PIDN:AAA60299.1; PID:g337752

R/Author: E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.

Biol. Chem. Hoppe-Seyler 369, 191-197, 1988

A/Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily

A/Reference number: S02489; MUID:89076505; PMID:3202964

A/Accession: S02489

A/Status: not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 21-141 <SA2>

R/Author: E.; Kim, H.S.; Smithies, O.; Maeda, N.

Gene 61, 329-338, 1987

A/Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three memt

A/Reference number: A91589; MUID:88185836; PMID:3446578

A/Accession: A29632



A/Note: the cystatin C gene isolated from the brain of an Icelandic patient with hereditary e)

R/Saitoh, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada, Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989

A/Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which

A/Reference number: A33400; MUID:89350945; PMID:2764935

A/Accession: A33400

A/Molecule type: DNA

A/Residues: 1-24, 'T', 26-146 <SAI>

A/Cross-references: GB:M27689; GB:M27690; NID:G181385; PIDN:AA52164.1; PID:G

R/Chieso, J.; Cowan, N.; Frangione, B

Biol. Chem. Hoppe-Seyler 369, 205-208, 1988

A/Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron

A/Reference number: S02751; MUID:89076507; PMID:3264504

A/Accession: S02751

A/Molecule type: DNA

A/Residues: 82-119 <GH2>

A/Cross-references: EMBL:M27769

A/Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown

R/Grubb, A.; Lotberg, H.

Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982

A/Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in th

A/Reference number: A01270; MUID:82222268; PMID:6283552

A/Accession: A01270

A/Molecule type: protein

A/Residues: 27-131, 'S', 133-146 <GRU>

R/Chieso, J.; Jenson, O.; Frangione, B.

Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986

A/Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland t

A/Reference number: A25434; MUID:86206076; PMID:3517880

A/Accession: A25434

A/Molecule type: protein

A/Residues: 37-93, 'Q', 95-146 <GHI>

R/Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.

Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983

A/Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystati

A/Reference number: S01461; MUID:84110059; PMID:6662498

A/Accession: S12288

A/Molecule type: protein

A/Residues: 27-73 <TRU>

R/Brzin, J.; Popovic, T.; Turk, V.

Biochem. Biophys. Res. Commun. 118, 103-109, 1984

A/Title: Human cystatin, a new protein inhibitor of cysteine proteinases.

A/Reference number: A32732; MUID:84128015; PMID:6365094

A/Accession: A32732

A/Molecule type: protein

A/Residues: 27-76 <BRZ>

R/Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.

Scand. J. Clin. Lab. Invest. 50, 85-93, 1990

A/Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst

A/Reference number: A60552; MUID:90193615; PMID:2315647

A/Accession: A60552

A/Molecule type: protein

A/Residues: 27-49, 'XX', 52-64 <OLA>

A/Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d

A/defective gene is not present in CSF but is found instead in amyloid deposits

R/Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.

Biol. Chem. Hoppe-Seyler 371, 575-580, 1990

A/Title: Different forms of human cystatin C.

A/Reference number: S10607; MUID:91025625; PMID:2222856

A/Accession: S10607

A/Molecule type: protein

A/Residues: 27-53 <POP>

A/Experimental source: urine, kidney disease

A/Note: truncated forms with amino ends at positions 35 and 36 of the precursor were als

R/Grubb, A.; Lotberg, H.; Barrett, A.J.

FEBS Lett. 170, 370-374, 1984

A/Title: The disulfide bridges of human cystatin C (gamma-trace) and chicken cystatin.

A/Reference number: S01462

A/Contents: annotation; disulfide bonds

R/Berti, P.J.; Storer, A.C.

Biochem. J. 302, 411-416, 1994

A/Title: Local pH-dependent conformational changes leading to proteolytic susceptibility

A/Reference number: S55305; MUID:94379969; PMID:8092991

A/Accession: S55305

A/Status: preliminary

A/Molecule type: protein

A/Residues: 27-49, 106-146 <BER>

C/Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl

I Patient with certain autoimmune diseases.

C/Comment: This protein is an inhibitor of cysteine proteinases and may serve an importar

C/Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorr

C/Genetic:

A/Gene: GDB:CST3

A/Cross-references: GDB:119817; OMIM:105150

A/Map position: 20p11.2-20p11.2

A/Intons: 81/3; 119/3

C/Superfamily: cystatin; cystatin homology

C/Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-146/Product: cystatin C #status experimental <MAT>

F:35-146/Domain: cystatin homology <CYS>

F:81-85/Region: inhibitory #status predicted

F:29/Modified site: hydroxyproline (Pro) (partial) #status experimental

F:99-109,123-143/Disulfide bonds: #status experimental

Query Match 45.5%; Score 65; DB 1; Length 146;  
Best Local Similarity 41.7%; Pred. No. 0.021;  
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDPKHFPIFRLVKQROV 25  
DB 59 EYNKASNDMYHSPALQVPRKQI 82

RESULT 5

UDBO

cystatin - bovine

N/Alternate names: thiol proteinase inhibitor

C/Species: Bos primigenius taurus (cattle)

C/Date: 28-Feb-1986 #sequence\_revision 28-Feb-1986 #text\_change 06-Dec-1996

C/Accession: A01271

R/Hirado, M.; Tsunawawa, S.; Sakiyama, F.; Nimobe, M.; Fujii, S.

FEBS Lett. 186, 41-45, 1985

A/Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase int

A/Reference number: A01271; MUID:85231205; PMID:3891407

A/Accession: A01271

A/Molecule type: protein

A/Residues: 1-112 <HIR>

C/Superfamily: cystatin; cystatin homology

C/Keywords: colostrum; cysteine proteinase inhibitor

F:2-112/Domain: cystatin homology <CYS>

F:48-52/Region: inhibitory #status predicted

F:66-76,90-110/Disulfide bonds: #status predicted

Query Match 43.4%; Score 62; DB 1; Length 112;  
Best Local Similarity 41.7%; Pred. No. 0.043;  
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDPKHFPIFRLVKQROV 25  
DB 26 EFNKRSNDAYQSHVVRVPRKQV 49

RESULT 6

S10587

cystatin C - rat

C/Species: Rattus sp. (rat)

C/Date: 21-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 16-Jul-1999

C/Accession: S10587

R/Barnard, F.; Barnard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Billard, M.; Gauchu

Biol. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990

A/Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylat

A/Reference number: S10587; MUID:90380276; PMID:2400577

A/Accession: S10587

A/Status: preliminary

A>Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cytatatin  
A:Reference number: S01461; MUID:84110059; PMID:6662498  
A:Accession: S01461  
A:Molecule type: protein  
A:Residues: 24-139 <TUR>  
R:Amanstetel, A.; Brown, M.A.; Kembhavi, A.A.; Nicklin, M.J.H.; Sayers, C.A.; Sunter, D.C.  
Biochem. J. 211, 129-138, 1983  
A>Title: Cystatin, a protein inhibitor of cysteine proteinaes. Improved purification fr  
A:Reference number: A37514; MUID:83256421; PMID:6403085  
A:Contents: annotation; characterization of protein  
R:Grubbs, A.; Lofberg, H.; Barrett, A.J.  
FEBS Lett. 170, 370-374, 1984  
A>Title: The disulfide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
A:Reference number: S01432  
A:Contents: annotation; disulfide bonds  
R:Auerwald, E.A.; Naegler, D.K.; Schlitz, A.J.; Engh, R.A.; Genenger, G.; Machleidt, W.  
Eur. J. Biochem. 224, 407-415, 1994  
A>Title: Production, inhibitory activity, folding and conformational analysis of an N-te  
A:Reference number: S48159; MUID:95010016; PMID:7925354  
A:Accession: S48159  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 24-139 <AUE>  
R:Lisner, B.; Krieglstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.  
FEBS Lett. 248, 162-168, 1989  
A>Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.  
A:Reference number: S04008; MUID:89252033; PMID:2721673  
A:Accession: S04008  
A:Molecule type: protein  
A:Residues: 97-114 <LAB>  
R:Cotella, R.; Bird, T.W.C.  
Gene 130, 175-181, 1993  
A>Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping t  
A:Reference number: JN0789; MUID:93366172; PMID:8359684  
A:Accession: JN0789  
A:Molecule type: DNA  
A:Residues: 1-139 <CO2>  
A:Cross-references: GS:M95725  
A>Note: authors failed to translate the codon for residue 115-Tyr  
C:Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases  
C:Genetics:  
A:Gene: Cam  
A:Introns: 76/3; 114/3  
C:Superfamily: cystatin; cystatin homology  
C:Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein  
F:1-73/Domain: signal sequence #status predicted <SIG>  
F:24-139/Product: cystatin, long form #status experimental <CLIF>  
E:30-139/Domain: cystatin homology <CYS>  
F:32-139/Product: cystatin, short form #status experimental <CYSP>  
F:76-80/Region: inhibitory #status predicted  
F:94-104,118-138/Disulfide bonds: #status experimental  
F:103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 46.2%; Score 66; DB 1; Length 139;  
Best Local Similarity 45.8%; Pred. No. 0.04;  
Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Cy 2 QYNKESDDKYHFRIFRVLYKVOROV 25  
Db 54 EYNPASNDKISSRRVRIASAKQL 77

RESULT 3  
B29632  
CyStatin SA precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1989 #sequence\_revision 30-Jun-1989 #text\_change 16-Jul-1999  
C:Accession: B29632; S02490; A41422; B27015  
R:Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.  
Gene 61, 329-338, 1987  
A>Title: Human cyteine-proteinase inhibitors: nucleotide sequence analysis of three mem  
A:Reference number: A31589; MUID:86185836; PMID:3446578  
A:Accession: B29632

A:Molecule type: DNA  
A:Residues: 1-141 <SA1>  
A:Cross-references: GB:M19673; GB:M19170; NID:9186403; PIDN:AAA6116.1; PID:9386826  
A:Note: The authors translated the codon GAC for residue 129 as Asn  
Biol. Chem. Hoppe-Seyler 369, 191-197, 1988  
A:Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily  
A:Reference number: S02489; MUID:89076505; PMID:3202964  
A:Accession: S02490  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 21-141 <SA2>  
J. Biochem. 102, 693-704, 1987  
A:Title: Characterization and amino acid sequence of a new acidic cysteine proteinase from  
A:Reference number: A41422; MUID:88139220; PMID:3436950  
A:Accession: A41422  
A:Molecule type: protein  
A:Residues: 25-141 <IS>  
R. Isenmura, S.; Saitoh, E.; Sanada, K.; Isenmura, M.; Ito, S.  
in Cysteine Proteinases and Their Inhibitors, Turk, V., ed., pp.497-505, Walter de Gruyter  
A:Title: Cystatin S and the related cysteine proteinase inhibitors in human saliva.  
A:Reference number: A27015  
A:Accession: B27015  
A:Molecule type: protein  
A:Residues: 25-134, 'D', 136-141 <IS2>  
C:Genetics:  
A:Gene: GDB: CST2  
A:Cross-references: GDB:119816; OMIM:123856  
A:Map position: 20p11.2-20p11.2  
C:Superfamily: cystatin; cystatin homology <CYS>  
P.30-141/Domain: Cystatin homology <CYS>

QY 2 OYNKESDDKYHRIPIVLTAKVOROV 25  
: ||| : ||| : ||| : |||  
Db 54 EYNKATEDEYRRLRVLPAREOI 77

RESULT 4

UDHU

Cystatin C precursor [validated] - human

M.Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma I

C.Species: Homo sapiens (man)

C.Date: 06-Jul-1982 #sequence\_revision 31-Mar-1991 #text\_change 08-Dec-2000

A.Accession: S10216; S000004; JI00095; A33440; S02751; A01570; A25434; S12288; A32732; A601

R.Abrahamson, M.; Olafsson, I.; Paladottir, A.; Ulvsbeck, M.; Lundwall, A.; Jensen, O.,

Biochem. J. 268, 287-294, 1990

Article: Structure and expression of the human cystatin C gene.

A.Reference number: S10216; PMID:90303202; PMID:2363674

A.Accession: S10216

A.Molecule type: DNA

A.Residues: 1-146 <AB1>

A.Cross-references: EMBL:X52255; NID:g30257; PIDD:CMA36497.1; PID:g296643

R.Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.

FBS Lett. 216, 229-233, 1987

Article: Molecular cloning and sequence analysis of cDNA coding for the precursor of the

A.Reference number: S00004; PMID:87219149; PMID:3395457

A.Accession: S00004

A.Molecule type: mRNA

A.Residues: 1-146 <AB2>

A.Cross-references: EMBL:X05607; NID:g30371; PIDD:CMA29096.1; PID:g755738

R.Levy, B.; Lopez-Otin, C.; Ghiso, J.; Gellner, D.; Fangerone, B.

J. Exp. Med. 169, 1771-1778, 1989

Article: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a

A.Reference number: JI0095; PMID:89235594; PMID:2541223

A.Accession: JI0095

A.Molecule type: DNA

A.Residues: 1-146 <LEV>

A.Cross-references: GB:X61681; NID:g30367; PIDD:CMA43856.2; PID:g4490944

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 5.19643 Seconds  
(without alignments)  
499.799 Million cell updates/sec

Title: US-09-941-314-9  
Perfect score: 143  
Sequence: 1 DQYNKESDDKYHFRFVRLKVKORQVTD 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	76	53.1	139	2	A45361	cystatin-related e
2	66	46.2	139	1	UDCH	cystatin precursor
3	66	46.2	141	2	B29632	cystatin SA precu
4	65	45.5	146	1	UDHU	cystatin C precu
5	62	43.4	112	1	UDBO	cystatin - bovine
6	62	43.4	120	2	S10587	cystatin C - rat
7	62	43.4	127	2	S07085	cystatin C precu
8	62	43.4	140	2	A36163	cystatin C precu
9	59	41.3	139	2	E64005	hypothetical prote
10	58	40.6	47	2	E62326	cystatin C, parot
11	56	39.2	162	2	A43428	oncho-cystatin - ne
12	54	37.8	141	1	UDHUP2	cystatin S precu
13	54	37.8	141	1	UDHUP2	cystatin SN precu
14	53	37.1	141	2	J01470	cystatin S precu
15	51	35.7	111	2	A28793	cystatin - puff ad
16	50.5	35.3	394	2	T33641	hypothetical prote
17	50	35.0	319	2	A86650	riamosyltransfera
18	50	35.0	393	2	D86168	hypothetical prote
19	49	34.3	442	2	T06640	hypothetical prote
20	49	34.3	498	2	T31871	hypothetical prote
21	49	34.3	525	1	KGHUGH	histidine-rich gly
22	49	34.3	1146	2	S46837	hypothetical prote
23	48	33.6	917	1	RDMUNH	nitrate reductase
24	47.5	33.2	505	1	PSXR4	outer capsid prote
25	47	32.9	224	2	SS4519	HLA1 protein - yea
26	47	32.9	370	2	T47762	hypothetical prote
27	47	32.9	602	2	S69198	prostaglandin G/H
28	47	32.9	602	2	S39782	cyclooxygenase 1 -
29	47	32.9	602	2	A35564	prostaglandin-endo

30	47	32.9	681	2	T01469	hypothetical prote
31	47	32.9	1973	2	G89608	protein B0272.5 [1
32	47	32.9	1973	2	T18686	hypothetical prote
33	47	32.9	2305	2	T15571	hypothetical prote
34	47	32.9	2305	2	B89608	protein C23P12.1 [
35	46	32.2	533	2	S57904	virR49 protein - S
36	46	32.2	533	2	UC5040	positive regulator
37	46	32.2	656	2	TS2064	dnaU-like protein
38	46	32.2	734	2	S37998	probable serine/th
39	45.5	31.8	160	2	E82710	conserved hypothet
40	45.5	31.8	262	2	S58288	heterocyst differe
41	45.5	31.8	273	2	T17629	hypothetical prote
42	45.5	31.8	430	2	B69620	phosphopyruvate hy
43	45.5	31.8	916	2	AG1902	acetyl-CoA synchet
44	45	31.5	73	2	J01921	hypothetical 8.5k
45	45	31.5	111	1	JC2040	cystatin - chum sa

## ALIGNMENTS

RESULT 1  
A45361  
cystatin-related epididymal specific protein - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C/Accession: A45361  
R/Cornwall, G.A.; Orgebin-Crist, M.C.; Hamu, S.R.  
Mol. Endocrinol. 6, 1653-1664, 1992  
A/Title: The CREB gene: a unique testis-regulated gene related to the cystatin family is  
A/Reference number: A45361; MUID:93078799; PMID:1280328  
A/Accession: A45361  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: nucleic acid  
A/Residues: 1-139 <COR>  
A/Cross-references: GB:S49926; NID:9260492; PIDN:AAC35390.1; PID:9260493  
A/Note: sequence extracted from NCBI backbone (NCBIP:118813)  
C/Superfamily: cystatin; cystatin homology  
F/28-139/Domain: cystatin homology <CYS>

Query Match 53.1%; Score 76; DB 2; Length 139;  
Best Local Similarity 50.0%; Pred. No. 0.00044;  
Matches 13; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 2 DQYNKESDDKYHFRFVRLKVKORQVTD 27  
Db 52 EYNKESDDKYVFLVDKILHAKLOITD 77

## RESULT 2

UDCH  
cystatin precursor - chicken  
N/Alternate names: cystatin 1; cysteine proteinase inhibitor; egg-white cystatin  
C/Species: Gallus gallus (chicken)  
C/Date: 03-Aug-1984 #sequence\_revision 12-Apr-1996 #text\_change 29-Oct-1999  
C/Accession: A4456; A01274; S01461; S48159; S04008; JN0789  
R/Colletta, R.; Sakaguchi, Y.; Nagase, H.; Bird, J.W.C.  
J. Biol. Chem. 264, 17154-17169, 1989  
A/Title: Chicken egg white cystatin. Molecular cloning, nucleotide sequence, and tissue c  
A/Reference number: A4456; MUID:9008873; PMID:2793849  
A/Accession: A4456  
A/Molecule type: mRNA  
A/Residues: 1-139 <COL>  
A/Cross-references: GB:U05077; NID:9211714; PIDN:AAA48744.1; PID:9211715  
R/Schwabe, C.; Anastasi, A.; Crow, H.; McDonald, J.K.; Barrett, A.U.  
Biochem. J. 217, 813-817, 1984  
A/Title: Cystatin. Amino acid sequence and possible secondary structure.  
A/Reference number: A01274; MUID:84178305; PMID:6712597  
A/Accession: A01274  
A/Molecule type: protein  
A/Residues: 24-139 <SCH>  
R/Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983

**THIS PAGE BLANK (USPTO)**



PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 48936  
LENGTH: 50  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL096677.18  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96  
OTHER INFORMATION: EST HUMAN HIT: A120857.1, EVALU 5.00e-23  
OTHER INFORMATION: SWISSPROT HIT: O60676, EVALU 1.00e-01  
US-09-864-761-48936

Query Match 89.5%; Score 128; DB 9; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1e-11;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQNKESDDKYHFRIFVLKVQROVD 24  
Db 27 DQNKESDDKYHFRIFVLKVQROVD 50

RESULT 10  
US-09-941-314-10  
Sequence 10, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 46  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-10

Query Match 87.4%; Score 125; DB 9; Length 46;  
Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NKESDDKYHFRIFVLKVQROVD 27  
Db 1 NKESDDKYHFRIFVLKVQROVD 24

RESULT 11  
US-09-941-314-14  
Sequence 14, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314

CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 52  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-14

Query Match 83.2%; Score 119; DB 9; Length 52;  
Best Local Similarity 100.0%; Pred. No. 2.3e-10;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KESDDKYHFRIFVLKVQROVD 27  
Db 1 KESDDKYHFRIFVLKVQROVD 23

RESULT 12  
US-10-264-049-2608  
Sequence 2608, Application US/10264049  
Publication No. US20040005579A1  
GENERAL INFORMATION:  
APPLICANT: Birse et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: P113P1  
CURRENT APPLICATION NUMBER: US/10/264,049  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/18569  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: US 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: PatentIn Ver. 3.1  
SEQ ID NO 2608  
LENGTH: 181  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-264-049-2608

Query Match 51.7%; Score 74; DB 15; Length 181;  
Best Local Similarity 54.2%; Pred. No. 0.004;  
Matches 13; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRIFVLKVQROVD 25  
Db 94 EYNKASNDAYHRAVVRVVKQV 117

RESULT 13  
US-09-775-932-16  
Sequence 16, Application US/09775932  
Patent No. US20020137671A1  
GENERAL INFORMATION:  
APPLICANT: University of British Columbia  
TITLE OF INVENTION: Production and use of Modified Cystatins  
FILE REFERENCE: 58069  
CURRENT APPLICATION NUMBER: US/09/775,932  
CURRENT FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: CA99/00717  
PRIOR FILING DATE: 1999-08-05  
PRIOR APPLICATION NUMBER: 60/095,503  
PRIOR FILING DATE: 1998-08-05  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 116  
TYPE: PRT  
ORGANISM: Gallus sp.  
US-09-775-932-16

GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-B and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-15

Query Match 95.8%; Score 137; DB 9; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8e-13;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKQVROVTD 27  
DB 1 QYNKESDDKXHFRIFRVLKQVROVTD 26

RESULT 8  
US-09-864-761-34822  
Sequence 34822, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/233,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 34822  
LENGTH: 50  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL109954.10  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2  
OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23  
OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01  
US-09-864-761-34822

Query Match 89.5%; Score 128; DB 9; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1e-11;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFRVLKQVRO 24  
DB 27 DQYNKESDDKXHFRIFRVLKQVRO 50

RESULT 9  
US-09-864-761-48936  
Sequence 48936, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/233,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408

```

? FILE REFERENCE: 00-81PC
? CURRENT APPLICATION NUMBER: US/09/941,314
? CURRENT FILING DATE: 2001-08-29
? PRIOR APPLICATION NUMBER: 60/230,230
? PRIOR FILING DATE: 2001-09-01
? NUMBER OF SEQ ID NOS: 19
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 8
? LENGTH: 35
? TYPE: PRT
? ORGANISM: Homo sapiens
? US-09-941-314-8

```

Query Match	100.0%	Score 143	DB 9	length 35
Best Local Similarity	100.0%	Pred. No. 4	2e-14	
Matches 27, Conservative	0	Mismatches	0	Gaps 0

[illegible]

RESULT 3  
 US-09-941-314-13  
 Sequence 13, Application US/09941314  
 Patent No. US20020142396A1  
 GENERAL INFORMATION:  
 APPLICANT: Zymogenetics, Inc.  
 TITLE OF INVENTION: Mammalian Cytactin-8  
 TITLE OF INVENTION: Inhibit Cancer Proco  
 FILE REFERENCE: 00-81PC  
 CURRENT APPLICATION NUMBER: US/09/941,314  
 CURRENT FILING DATE: 2001-08-29  
 PRIOR APPLICATION NUMBER: 60/230,230  
 PRIOR FILING DATE: 2001-09-01  
 NUMBER OF SEQ ID NOS: 19  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO 13  
 LENGTH: 49  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-941-314-13

Query March	100.0%	Score 143;	DB 9;	Length 49;
Best Local Similarity	100.0%	Pred. No. 6.1e-14;		
Matches 27;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	Db
1 DQYNKESDDKYHFRIFRLVKVQRQVTD 27	1 DQYNKESDDKYHFRIFRLVKVQRQVTD 27

```

RESULT 4
US-09-941-314-3
; Sequence 3, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: Zymogenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIORITY FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ. ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO 3
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-3

```

Query Match	100.0%	Score 143;	DB 9;	Length 115;
Best local similarity	100.0%;	pred. No. 1	5e-13;	
Matches	27;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

**OY**      1 DQYNKESDDKHFRIFRVLKVQRVTD 27  
         |||||  
**Db**     30 DQYNKESDDKHFRIFRVLKVQRVTD 56

RESULT 5  
US-09-941-314-4  
; Sequence 4, Application US/09941314  
; Patent No. US20020142396A1  
GENERAL INFORMATION

```

? APPLICANT: ZymoGenetics, Inc.
? TITLE OF INVENTION: Mammalian Cystatin-B and Its Use to
? TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
? FILE REFERENCE: 00-81PC
? CURRENT APPLICATION NUMBER: US/09/941,314
? CURRENT FILING DATE: 2001-08-29
? PRIOR APPLICATION NUMBER: 60/230,220
? PRIOR FILING DATE: 2001-09-01
? NUMBER OF SEQ ID NOS: 19
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 4
? LENGTH: 117
? TYPE: PRT
? ORGANISM: Homo sapiens
? US-09-941-4

```

Query March	100.0%	Score 143;	DB 9;	Length 117;
Best Local Similarity	100.0%	Pred. No. 1.6e-13;		
Matches 27; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 DQYNKESDDKCHFRIFRVLKQVRQVTD 27  
Db 32 DQYNKESDDKCHFRIFRVLKQVRQVTD 58

```

RESULT 6
US-09-941-314-2
Sequence 2, Application US/09941314
Patent No. US200200142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-B
TITLE OF INVENTION: Inhibit Cancer Proce
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 137
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-2

```

Query Match	100.0%	Score 143;	DB 9;	Length 137;
Best Local Similarity	100.0%	Pred. No. 1.8e-13;		
Matches 27;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 DQYNKESDDKKHFRI FRVLKVQRQVTD 27  
|||  
Db 52 DQYNKESDDKKHFRI FRVLKVQRQVTD 78

RESULT 7  
US-09-941-314-15  
; Sequence 15, Application US/09941314  
; Patent No. US20020142396A1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 16.3393 Seconds  
(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-9  
Perfect score: 143  
Sequence: 1 DQYKESDDKYHFRIRFLVKVQRQVTD 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
12: /cgn2\_6/prodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/prodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	143	100.0	27	9	US-09-941-314-9	Sequence 9, Appl1
2	143	100.0	35	9	US-09-941-314-8	Sequence 8, Appl1
3	143	100.0	49	9	US-09-941-314-13	Sequence 13, Appl1
4	143	100.0	115	9	US-09-941-314-3	Sequence 3, Appl1
5	143	100.0	117	9	US-09-941-314-4	Sequence 4, Appl1
6	143	100.0	137	9	US-09-941-314-2	Sequence 2, Appl1
7	137	95.8	80	9	US-09-941-314-15	Sequence 15, Appl1
8	128	89.5	50	9	US-09-864-761-34822	Sequence 34822, A
9	125	87.4	46	9	US-09-864-761-48936	Sequence 48936, A
10	119	83.2	52	9	US-09-941-314-10	Sequence 10, Appl1
11	119	83.2	52	9	US-09-941-314-14	Sequence 14, Appl1
12	74	51.7	181	15	US-10-264-049-2608	Sequence 2608, Ap
13	66	46.2	116	9	US-09-775-932-16	Sequence 16, Appl1
14	66	46.2	121	9	US-09-775-932-8	Sequence 8, Appl1
15	66	46.2	139	8	US-08-849-303-15	Sequence 15, Appl1

16	66	46.2	139	9	US-09-969-834-4	Sequence 4, Appl1
17	66	46.2	141	8	US-08-849-303-24	Sequence 24, Appl1
18	66	46.2	141	9	US-09-940-497-6	Sequence 6, Appl1
19	65	45.5	120	9	US-09-775-932-2	Sequence 2, Appl1
20	65	45.5	146	8	US-08-849-303-17	Sequence 17, Appl1
21	65	45.5	146	9	US-09-940-497-3	Sequence 3, Appl1
22	65	45.5	146	9	US-09-969-834-3	Sequence 3, Appl1
23	65	45.5	146	14	US-10-329-428-3	Sequence 3, Appl1
24	65	45.5	146	14	US-10-376-564-47	Sequence 47, Appl1
25	63	44.1	145	9	US-09-740-638-2	Sequence 2, Appl1
26	63	44.1	145	13	US-10-006-467-2	Sequence 2, Appl1
27	63	44.1	145	14	US-10-235-148-2	Sequence 2, Appl1
28	62	43.4	112	8	US-08-849-303-16	Sequence 16, Appl1
29	62	43.4	118	9	US-09-775-932-24	Sequence 24, Appl1
30	62	43.4	127	8	US-08-849-303-19	Sequence 19, Appl1
31	62	43.4	140	8	US-08-849-303-18	Sequence 18, Appl1
32	62	43.4	140	14	US-10-376-564-46	Sequence 46, Appl1
33	62	43.4	140	14	US-10-376-564-48	Sequence 48, Appl1
34	61	42.7	145	14	US-09-775-932-18	Sequence 18, Appl1
35	57	39.9	145	14	US-10-168-425-14	Sequence 14, Appl1
36	56	39.2	162	8	US-08-849-303-25	Sequence 25, Appl1
37	55	38.5	18	9	US-09-941-314-7	Sequence 7, Appl1
38	55	38.5	36	9	US-09-941-314-6	Sequence 6, Appl1
39	54	37.8	121	9	US-09-775-932-4	Sequence 4, Appl1
40	54	37.8	121	9	US-09-775-932-6	Sequence 6, Appl1
41	54	37.8	141	8	US-08-849-303-21	Sequence 21, Appl1
42	54	37.8	141	8	US-08-849-303-22	Sequence 22, Appl1
43	54	37.8	141	9	US-09-940-497-5	Sequence 5, Appl1
44	54	37.8	141	9	US-09-940-497-7	Sequence 7, Appl1
45	54	37.8	141	9	US-09-974-298-141	Sequence 141, Appl1

## ALIGNMENTS

RESULT 1  
US-09-941-314-9  
; Sequence 9, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: Zymogenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 27  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-9

Query Match 100.0%; Score 143; DB 9; Length 27;  
Best Local Similarity 100.0%; Pred. No. 3.2e-14;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQYKESDDKYHFRIRFLVKVQRQVTD 27  
Db 1 DQYKESDDKYHFRIRFLVKVQRQVTD 27

RESULT 2  
US-09-941-314-8  
; Sequence 8, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: Zymogenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein

**THIS PAGE BLANK (USPTO)**

APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: Human Cystatin E  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/241,376  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/461,030  
FILING DATE: 05-NOV-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/461,030  
FILING DATE: 05-JUN-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PP202PID1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301 309 8504  
TELEFAX: 301 309 8512  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Cystatin SN  
US-09-241-376-6

Query Match 46.2%; Score 66; DB 4; Length 141;  
Best Local Similarity 41.7%; Pred. No. 0.0045;  
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDXYHFRIFRYLKYQROV 25  
Db 54 EYNKATEDERYRRLRYLRAREQI 77

RESULT 14  
US-09-940-497-6  
Sequence 6, Application US/09940497  
Patent No. 6617132  
GENERAL INFORMATION:  
APPLICANT: Ni et al.  
TITLE OF INVENTION: Human Cystatin E  
FILE REFERENCE: PP202PID2  
CURRENT APPLICATION NUMBER: US/09/940,497  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: US 09/241,376  
PRIOR FILING DATE: 1999-02-02  
PRIOR APPLICATION NUMBER: US 08/744,138  
PRIOR FILING DATE: 1996-11-05  
PRIOR APPLICATION NUMBER: US 08/461,030  
PRIOR FILING DATE: 1995-06-05  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: Patent In Release #1.0, Version #1.30  
US-09-940-497-6

SEQ ID NO 6  
LENGTH: 141  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-940-497-6

Query Match 46.2%; Score 66; DB 4; Length 141;  
Best Local Similarity 41.7%; Pred. No. 0.0045;  
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDXYHFRIFRYLKYQROV 25  
Db 54 EYNKATEDERYRRLRYLRAREQI 77

## RESULT 15

US-08-849-303-24  
Sequence 24, Application US/08849303  
Patent No. 6680424  
GENERAL INFORMATION:  
APPLICANT: Atkinson, Howard J.  
ADDRESSEE: McPherson, Michael J.  
APPLICANT: Urwin, Peter E.  
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/849,303  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1321-1-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
US-08-849-303-24

Query Match 46.2%; Score 66; DB 4; Length 141;  
Best Local Similarity 41.7%; Pred. No. 0.0045;  
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDXYHFRIFRYLKYQROV 25  
Db 54 EYNKATEDERYRRLRYLRAREQI 77

Search completed: March 18, 2004, 14:25:36  
Job time: 6.91071 secs

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/791,522  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0193 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 118195  
US-09-314-777-4

Query Match 46.2%; Score 66; DB 3; Length 139;  
Best Local Similarity 45.8%; Pred. No. 0.0044;  
Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 2 QYNKESDDKXHFRIPLVLYKVOROV 25  
Db 54 EYNRASNDKXSRVVRVISAQRQL 77

RESULT 11  
US-08-849-303-15  
Sequence 15, Application US/08849303  
Patent No. 6680424  
GENERAL INFORMATION:  
APPLICANT: Atkinson, Howard J.  
APPLICANT: McPherson, Michael J.  
APPLICANT: Urvn, Peter E.  
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/849,303  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1321-1-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-849-303-15

Query Match 46.2%; Score 66; DB 4; Length 139;  
Best Local Similarity 45.8%; Pred. No. 0.0044;  
Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 2 QYNKESDDKXHFRIPLVLYKVOROV 25  
Db 54 EYNRASNDKXSRVVRVISAQRQL 77

RESULT 12  
US-08-744-138-6  
Sequence 6, Application US/08744138  
Patent No. 6011012  
GENERAL INFORMATION:  
APPLICANT: Gentz, Reiner L.  
APPLICANT: Ni, Jian  
APPLICANT: Rosen, Craig A.  
APPLICANT: Yu, Guo-liang  
TITLE OF INVENTION: Human Cystatin B  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/744,138  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF202P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301 309 8504  
TELEFAX: 301 309 8512  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 141 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Cystatin SN  
US-08-744-138-6

Query Match 46.2%; Score 66; DB 3; Length 141;  
Best Local Similarity 41.7%; Pred. No. 0.0045;  
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QYNKESDDKXHFRIPLVLYKVOROV 25  
Db 54 EYNKATEDYTRILRVLRARBOI 77

RESULT 13  
US-09-241-376-6  
Sequence 6, Application US/09241376  
Patent No. 6306477  
GENERAL INFORMATION:  
APPLICANT: Gentz, Reiner L.  
APPLICANT: Ni, Jian  
APPLICANT: Rosen, Craig A.

```
RESULT 7
US-09-775-932-16
; Sequence 16, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cycatans
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-775-932-16

Query Match          46.2%; Score 66; DB 4; Length 116;
Best Local Similarity 45.8%; Pred. No. 0.0036;
Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY      2 QYNKESDXYHFRIFRYLKVQROV 25
DB      31 EYNRASNDKYSRVVRVISAQRL 54

RESULT 8
US-09-775-932-8
; Sequence 8, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cycatans
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-775-932-8

Query Match          46.2%; Score 66; DB 4; Length 121;
Best Local Similarity 41.7%; Pred. No. 0.0037;
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY      2 QYNKESDXYHFRIFRYLKVQROV 25
DB      34 EYNKATEDEYRRLRLVLRARAEQI 57

RESULT 9
US-08-791-522-4
; Sequence 4, Application US/08791522
; Patent No. 5935817
; GENERAL INFORMATION:
; APPLICANT: Bardman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
```

```
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,522
FILING DATE: Filed Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0193 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 118195
US-08-791-522-4

Query Match          46.2%; Score 66; DB 2; Length 139;
Best Local Similarity 45.8%; Pred. No. 0.0044;
Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY      2 QYNKESDXYHFRIFRYLKVQROV 25
DB      54 EYNRASNDKYSRVVRVISAQRL 77

RESULT 10
US-09-314-777-4
; Sequence 4, Application US/09314777
; Patent No. 610686
; GENERAL INFORMATION:
; APPLICANT: Bardman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/314,777
; FILING DATE:
; CLASSIFICATION:
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 14:07:11 ; Search time 6.91071 Seconds  
(without alignments)  
201.701 Million cell updates/sec

Title: US-09-941-314-9  
Perfect score: 143  
Sequence: 1 DQYKESDDKXHFRLVKYQROVTD 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.dep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.dep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.dep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.dep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.dep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfile1.dep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	56.6	141	3	US-09-431-480-2
2	81	56.6	141	3	US-09-617-302-2
3	77	53.8	142	3	US-09-431-480-4
4	77	53.8	142	3	US-09-617-302-4
5	76	53.1	142	3	US-09-431-480-3
6	76	53.1	142	3	US-09-617-302-3
7	66	46.2	116	4	US-09-775-932-16
8	66	46.2	121	4	US-09-775-932-8
9	66	46.2	139	2	US-08-791-522-4
10	66	46.2	139	3	US-09-314-777-4
11	66	46.2	139	4	US-08-849-303-15
12	66	46.2	141	3	US-08-744-138-6
13	66	46.2	141	4	US-09-241-376-6
14	66	46.2	141	4	US-09-940-497-6
15	66	46.2	141	4	US-08-849-303-24
16	65	45.5	120	4	US-09-775-932-2
17	65	45.5	120	6	5432264-4
18	65	45.5	145	2	US-08-832-535-11
19	65	45.5	146	2	US-08-791-522-3
20	65	45.5	146	3	US-08-744-138-3
21	65	45.5	146	3	US-09-019-485-4
22	65	45.5	146	3	US-09-314-777-3
23	65	45.5	146	3	US-09-431-480-6
24	65	45.5	146	3	US-09-617-302-6
25	65	45.5	146	3	US-09-241-376-3
26	65	45.5	146	4	US-09-528-436B-3
27	65	45.5	146	4	US-09-886-319A-47

28	65	45.5	146	4	US-09-940-497-3	Sequence 3, Appli
29	65	45.5	146	4	US-09-976-594-37	Sequence 37, Appli
30	65	45.5	146	4	US-08-849-303-17	Sequence 17, Appli
31	65	45.5	146	5	PCT-US95-07135-9	Sequence 9, Appli
32	65	45.5	146	6	5432264-6	Patent No. 5432264
33	63	44.1	56	1	US-08-538-165A-4	Sequence 4, Appli
34	62	43.4	112	4	US-08-849-303-16	Sequence 16, Appli
35	62	43.4	118	4	US-09-775-932-24	Sequence 24, Appli
36	62	43.4	127	4	US-08-849-303-19	Sequence 19, Appli
37	62	43.4	140	3	US-09-431-480-5	Sequence 5, Appli
38	62	43.4	140	3	US-09-617-302-5	Sequence 5, Appli
39	62	43.4	140	4	US-09-886-319A-46	Sequence 46, Appli
40	62	43.4	140	4	US-09-886-319A-48	Sequence 48, Appli
41	62	43.4	140	4	US-08-849-303-18	Sequence 18, Appli
42	61	42.7	111	4	US-09-775-932-18	Sequence 18, Appli
43	56	39.2	162	4	US-08-849-303-25	Sequence 25, Appli
44	54	37.8	56	1	US-08-538-165A-1	Sequence 1, Appli
45	54	37.8	56	1	US-08-538-165A-7	Sequence 7, Appli

## ALIGNMENTS

```

RESULT 1
US-09-431-480-2
; Sequence 2, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; EARLIER FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-2

Query Match      56.6%; Score 81; DB 3; Length 141;
Best local Similarity 53.8%; Pred. No. 1.9e-05;
Matches 14; Conservative 6; Mismatches 6; Gaps 0;

Cy      2 DQYKESDDKXHFRLVKYQROVTD 27
Db      54 EYNKASMDLYNFRVVDLKSQEOITD 79

RESULT 2
US-09-617-302-2
; Sequence 2, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; EARLIER FILING DATE: 2000-07-17
; EARLIER APPLICATION NUMBER: 09/431,480
; EARLIER FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22

```

QY 1 DOYNKESDDKXHFRIFRVLKVRQ 24  
 |||  
 Db 27 DOYNKESDDKXHFRIFRVLKVRQ 50

## RESULT 15

AAMS4871  
 ID AAMS4871 standard; protein; 50 AA.

AC AAMS4871;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 26976.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;  
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX brains.

XX Example 4; SEQ ID NO 26976; 650bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX brain. They can be used to measure gene expression in brain cell samples,

XX which may enable the diagnosis and improved treatment of nervous system

XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX epilepsy and cancers. The present sequence is a protein encoded by one of

XX the probes of the invention

XX SQ Sequence 50 AA;

XX Query Match 89.5%; Score 128; DB 4; Length 50;

XX Best Local Similarity 100.0%; Pred. No. 2.9e-12;

XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DOYNKESDDKXHFRIFRVLKVRQ 24  
 |||  
 Db 27 DOYNKESDDKXHFRIFRVLKVRQ 50

Search completed: March 18, 2004, 14:14:59  
 Job time : 24.4821 secs



PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488897/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human placenta.  
XX  
PS Claim 27; SEQ ID NO 27814; 654pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes (SENP;  
CC see A13315-A157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders  
XX  
SQ Sequence 50 AA;  
XX  
Query Match 89.5%; Score 128; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 2.9e-12;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 DOYNKESDDKYHPRIFRVLKQVQRQ 24  
Db 27 DOYNKESDDKYHPRIFRVLKQVQRQ 50  
XX  
RESULT 11  
ABB32389  
ID ABB32389; standard; peptide; 50 AA.  
XX  
AC ABB32389;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Peptide #5040 encoded by breast cell single exon nucleic acid probe.  
XX  
XX Human; microarray; single exon probe; gene expression; breast; disease;  
KM cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157271-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000662.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-496933/54.  
XX  
XX New spatially-addressable set of single exon nucleic acid probes, useful  
PT for measuring gene expression in sample derived from human breast,  
PT comprises number of single exon nucleic acid probes.  
XX  
PS Claim 27; SEQ ID NO 15357; 327pp + Sequence listing; English.

XX  
CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting the  
CC probes with a collection of detectably labelled nucleic acids derived  
CC from mRNA of human breast, and then measuring the label bound to each  
CC probe of the microarray. The probes are useful for verifying the  
CC expression of regions of genomic DNA predicted to encode proteins. They  
CC are useful for gene discovery, and for determining predisposition and/or  
CC prognosing breast disease. Gene expression analysis is useful for  
CC assessing the toxicity of chemical agents on cells. The microarray of  
CC this invention presents a far greater diversity of probes for measuring  
CC gene expression, with far less bias than expressed sequence tag  
CC microarrays. The method is suitable for rapid production of functional  
CC information from genomic sequence. The present sequence is a peptide  
CC encoded by a single exon nucleic acid probe of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 50 AA;  
XX  
Query Match 89.5%; Score 128; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 2.9e-12;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 DOYNKESDDKYHPRIFRVLKQVQRQ 24  
Db 27 DOYNKESDDKYHPRIFRVLKQVQRQ 50  
XX  
RESULT 12  
ABB28913  
ID ABB28913; standard; peptide; 50 AA.  
XX  
AC ABB28913;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Peptide #1564 encoded by breast cell single exon nucleic acid probe.  
XX  
XX Human; microarray; single exon probe; gene expression; breast; disease;  
KM cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157271-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000662.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-496933/54.  
XX  
XX New spatially-addressable set of single exon nucleic acid probes, useful  
PT for measuring gene expression in sample derived from human breast,  
PT comprises number of single exon nucleic acid probes.  
XX  
PS Claim 27; SEQ ID NO 11881; 327pp + Sequence listing; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon

```

RESULT 8
AAM15096
ID AAM15096 standard; protein; 50 AA.
XX
AC AAM15096;
XX
DT 12-OCT-2001 (first entry)
DE Peptide #1530 encoded by probe for measuring cervical gene expression.
KW probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
FN WO200157278-A2.
PD
ID 09-AUG-2001.
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
XX
XX Claim 27; SEQ ID NO 19922; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SNP: see A110068-A128459). The present sequence is a peptide encoded
XX by one such probe. The SNPs are derived from human HeLa cells. The SNPs
XX can be used to produce a single exon microarray, which can be used for
XX measuring human gene expression in a sample derived from human cervical
XX epithelial cells. By measuring gene expression, the probes are therefore
XX useful in grading and/or staging of diseases of the cervix, notably
XX cervical cancer. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences
XX
SQ Sequence 50 AA;
XX
XX Query Match 89.5%; Score 128; DB 4; Length 50;
XX Best Local Similarity 100.0%; Pred. NO. 2.9e-12;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 DQYNKESDDKTHFRIFRLVLYKQQRQ 24
XX |||||||||||||||||||
XX |||||||||||||||||||
DB 27 DQYNKESDDKTHFRIFRLVLYKQQRQ 50
XX
RESULT 9
AAB34086
ID AAB34086 standard; peptide; 50 AA.
XX
AC AAB34086;
XX
DT 04-FEB-2002 (first entry)
DE Peptide #1592 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.

```

```

XX Homo sapiens.
OS
XX MO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
XX
XX Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human fetal liver. The
XX present sequence is a peptide encoded by a single exon nucleic acid probe
XX of the invention. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 50 AA;
XX
XX
XX Query Match 89.5%; Score 128; DB 4; Length 50;
XX Best Local Similarity 100.0%; Pred. No. 2.9e-12;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 DQYKESDDKKYFRIFRVLKVQRQ 24
XX ||||||||||||||||||||
XX 27 DQYKESDDKKYFRIFRVLKVQRQ 50
XX
XX
XX RESULT 10
XX AAM27545
XX ID AAM27545 standard; protein; 50 AA.
XX
XX AAM27545;
XX
XX 17-OCT-2001 (first entry)
XX
XX Peptide #1582 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
XX Homo sapiens.
XX
XX MO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX

```

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQYNKESDDKXHFRIFRVLKVGROVTD 27  
Db 32 DQYNKESDDKXHFRIFRVLKVGROVTD 58

## RESULT 6

AAU79852  
ID AAU79852 standard; protein; 137 AA.

AC AAU79852;

XX 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8).

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KW sperm motility; fertilisation.

OS Homo sapiens.

PN WO200220567-A2.

XX 14-MAR-2002.

PF 29-AUG-2001; 2001WO-US026668.

XX 01-SEP-2000; 2000US-0230230P.

PA (ZYMO ) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting

PS spermatoogenesis, and inhibiting cancer procoagulant protein which leads

XX Claim 2; Page 93-94; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)

CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant

CC protein in an individual and thus inhibiting the thrombotic events

CC associated with cancer; promoting spermatogenesis, modulating seminal

CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm

CC motility and fertilisation; and as antigenic peptides to generate

CC antibodies. Zcys8 is useful as research reagent for characterising sites

CC of interaction between Zcys8 and its receptor. Zcys8 is useful in

CC enhancing fertilisation during assisted reproduction in humans and in

CC animals. Anti-(I) antibodies are useful to screen biological samples like

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

CC presence of Zcys8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the

CC expression of a Zcys8 gene in a biological sample and Zcys8

CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (I) is useful in determining whether a subject's

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene

CC copy number changes, insertions, deletions, restriction site changes and

CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This is the amino acid sequence of human cystatin-8 (Zcys8)

XX Sequence 137 AA;

Query Match 100.0%; Score 143; DB 5; Length 137;

Best Local Similarity 100.0%; Pred. No. 4,4e-14;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DQYNKESDDKXHFRIFRVLKVGROVTD 27

Db 52 DQYNKESDDKXHFRIFRVLKVGROVTD 78

Qy 2 DQYNKESDDKXHFRIFRVLKVGROVTD 27  
Db 1 DQYNKESDDKXHFRIFRVLKVGROVTD 26

## RESULT 7

AAU79865  
ID AAU79865 standard; peptide; 80 AA.

AC AAU79865;

XX 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #13.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KW sperm motility; fertilisation; antigenic peptide.

OS Homo sapiens.

PN WO200220567-A2.

XX 14-MAR-2002.

PF 29-AUG-2001; 2001WO-US026668.

XX 01-SEP-2000; 2000US-0230230P.

PA (ZYMO ) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting

PS spermatoogenesis, and inhibiting cancer procoagulant protein which leads

XX Claim 2; Page 98; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)

CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant

CC protein in an individual and thus inhibiting the thrombotic events

CC associated with cancer; promoting spermatogenesis, modulating seminal

CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm

CC motility and fertilisation; and as antigenic peptides to generate

CC antibodies. Zcys8 is useful as research reagent for characterising sites

CC of interaction between Zcys8 and its receptor. Zcys8 is useful in

CC enhancing fertilisation during assisted reproduction in humans and in

CC animals. Anti-(I) antibodies are useful to screen biological samples like

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

CC presence of Zcys8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the

CC expression of a Zcys8 gene in a biological sample and Zcys8

CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (I) is useful in determining whether a subject's

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene

CC copy number changes, insertions, deletions, restriction site changes and

CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX Sequence 80 AA;

Query Match 95.8%; Score 137; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 2e-13;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 DQYNKESDDKXHFRIFRVLKVGROVTD 27

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
SQ Sequence 49 AA;

Query Match 100.0%; Score 143; DB 5; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.4e-14;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFRVLKVKORQVTD 27  
1 DQYNKESDDKXHFRIFRVLKVKORQVTD 27  
DB 1 DQYNKESDDKXHFRIFRVLKVKORQVTD 27

RESULT 4  
AAU79853  
ID AAU79853 standard; protein; 115 AA.

XX AAU79853;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #1.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic fragment.

XX Homo sapiens.

OS WO200220567-A2.

PN 14-MAR-2002.

PD 29-AUG-2001; 2001WO-US026868.

PF 01-SEP-2000; 2000US-0230230P.

PR (ZYMO) ZYMOGENETICS INC.

PA Holloway JL, Gao Z, Bishop PD;

PI WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 94; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic fragment of human cystatin-8  
CC (Zcys8)

SQ Sequence 115 AA;

Query Match 100.0%; Score 143; DB 5; Length 115;  
Best Local Similarity 100.0%; Pred. No. 3.6e-14;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFRVLKVKORQVTD 27  
1 DQYNKESDDKXHFRIFRVLKVKORQVTD 27  
DB 30 DQYNKESDDKXHFRIFRVLKVKORQVTD 56

RESULT 5  
AAU79854  
ID AAU79854 standard; protein; 117 AA.

XX AAU79854;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #2.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic fragment.

XX Homo sapiens.

OS WO200220567-A2.

PN 14-MAR-2002.

PD 29-AUG-2001; 2001WO-US026868.

PF 01-SEP-2000; 2000US-0230230P.

PR (ZYMO) ZYMOGENETICS INC.

PA Holloway JL, Gao Z, Bishop PD;

PI WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 94-95; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic fragment of human cystatin-8  
CC (Zcys8)

SQ Sequence 117 AA;

Query Match 100.0%; Score 143; DB 5; Length 117;  
Best Local Similarity 100.0%; Pred. No. 3.7e-14;

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (1) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (1) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX Sequence 27 AA:

Query Match 100.0%; Score 143; DB 5; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-15;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DOYNKESDDKXHFRIFRVLKVRQVTD 27  
 Db 1 DOYNKESDDKXHFRIFRVLKVRQVTD 27

RESULT 2  
 ID AAU79858 standard; peptide; 35 AA.  
 AC AAU79858;  
 XX  
 XX 15-JUL-2002 (first entry)  
 DT  
 XX Human cystatin-8 (Zcys8) antigenic fragment #6.  
 DE  
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic peptide.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200220567-A2.  
 PN  
 XX 14-MAR-2002.  
 PD  
 XX 29-AUG-2001; 2001MO-US026868.  
 PF  
 XX 01-SEP-2000; 2000US-0230230P.  
 PR  
 XX (ZYMO ) ZYMOGENETICS INC.  
 PA  
 XX Holloway JL, Gao Z, Bishop PD;  
 PI  
 XX WPI; 2002-383044/41.  
 DR  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 PS  
 XX Claim 2; Page 96; 100pp; English.  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(1) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (1) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (1) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX Sequence 35 AA:

Query Match 100.0%; Score 143; DB 5; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-15;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DOYNKESDDKXHFRIFRVLKVRQVTD 27  
 Db 9 DOYNKESDDKXHFRIFRVLKVRQVTD 35

RESULT 3  
 ID AAU79863 standard; peptide; 49 AA.  
 AC AAU79863;  
 XX  
 XX 15-JUL-2002 (first entry)  
 DT  
 XX Human cystatin-8 (Zcys8) antigenic fragment #11.  
 DE  
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic peptide.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200220567-A2.  
 PN  
 XX 14-MAR-2002.  
 PD  
 XX 29-AUG-2001; 2001MO-US026868.  
 PF  
 XX 01-SEP-2000; 2000US-0230230P.  
 PR  
 XX (ZYMO ) ZYMOGENETICS INC.  
 PA  
 XX Holloway JL, Gao Z, Bishop PD;  
 PI  
 XX WPI; 2002-383044/41.  
 DR  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 PS  
 XX Claim 2; Page 97-98; 100pp; English.  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(1) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (1) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (1) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using SW model

Run on: March 18, 2004, 14:02:15 ; Search time 24.4821 Seconds  
(without alignments)  
311.606 Million cell updates/sec

Title: US-09-941-314-9

Sequence: 1 DQYNKESDDKYHFRIFRVLKVRQVTD 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: geneseqp1980a:\*
- 2: geneseqp1980a:\*
- 3: geneseqp2000a:\*
- 4: geneseqp2000a:\*
- 5: geneseqp2002a:\*
- 6: geneseqp2003a:\*
- 7: geneseqp2003b:\*
- 8: geneseqp2004a:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	27	5	Aau79859 Human cys
2	143	100.0	35	5	Aau79858 Human cys
3	143	100.0	49	5	Aau79863 Human cys
4	143	100.0	115	5	Aau79853 Human cys
5	143	100.0	117	5	Aau79854 Human cys
6	143	100.0	137	5	Aau79852 Human cys
7	137	95.8	80	5	Aau79865 Human cys
8	128	89.5	50	4	Aau79865 Peptide #
9	128	89.5	50	4	ABH34086 Peptide #
10	128	89.5	50	4	ABH34086 Peptide #
11	128	89.5	50	4	ABH34086 Peptide #
12	128	89.5	50	4	ABH34086 Peptide #
13	128	89.5	50	4	ABH34086 Peptide #
14	128	89.5	50	4	ABH34086 Peptide #
15	128	89.5	50	4	ABH34086 Peptide #
16	128	89.5	50	4	ABH34086 Peptide #
17	128	89.5	50	4	ABH34086 Peptide #
18	128	89.5	50	4	ABH34086 Peptide #
19	128	89.5	50	4	ABH34086 Peptide #
20	119	83.2	52	5	Aau79860 Human cys
21	81	56.6	141	3	AAV96576 Murine cys
22	81	56.6	141	4	AAE02403 Murine cys
23	81	56.6	141	4	AAE04432 Murine cys
24	77	53.8	92	2	AAW78259 Fragment
25	77	53.8	113	6	ADA57563 Human sec

26	77	53.8	113	6	ADA41457 Human sec
27	77	53.8	113	7	ADG74577 Human sec
28	77	53.8	113	7	ADD38088 Human sec
29	77	53.8	114	2	AAW78153 Human sec
30	77	53.8	123	2	AAW78260 Fragment
31	77	53.8	142	2	AAW78258 Fragment
32	77	53.8	142	4	AAE02405 Human cys
33	77	53.8	142	4	AAE04434 Human cys
34	77	53.8	142	6	ADA57231 Human sec
35	77	53.8	142	6	ADA41112 Human sec
36	77	53.8	142	7	ADC74335 Human sec
37	77	53.8	142	7	ADD37980 Human sec
38	77	53.8	142	7	ADD46706 Human pro
39	77	53.8	142	7	ADD46710 Human pro
40	76	53.1	142	4	AAE02404 Murine cys
41	76	53.1	142	4	AAE04433 Mouse cys
42	76	53.1	143	6	ADA14374 Mouse spe
43	74	51.7	181	5	ABP41476 Human ova
44	67	46.9	116	3	AAE01203 Human mut
45	67	46.9	121	3	AAE01165 Human mut

## ALIGNMENTS

RESULT 1  
AAU79859 standard; peptide; 27 AA.  
XX  
AC AAU79859;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human cystatin-8 (Zcys8) antigenic fragment #7.  
XX  
KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic peptide.  
XX  
OS Homo sapiens.  
XX  
FN WO200220567-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US026868.  
XX  
PR 01-SEP-2000; 2000US-0230230P.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Holloway JL, Gao Z, Bishop PD;  
XX  
DR WPI; 2002-383044/41.  
XX  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX  
PS Claim 2; Page 96; 100pp; English.  
XX  
The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterizing sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large

**THIS PAGE BLANK (USPTO)**

Qy 7 ITDQYNKESDDKXHFRIFRVLKVGROV 33  
 Db 68 VLTFRVNOQSNDEYHLMPIKLVSSQV 94

## RESULT 13

ID 080125 PRELIMINARY; PRT; 464 AA.  
 AC 080125;  
 DT 01-JUN-2003 (TEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)  
 DE 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Refun Long form.  
 OS Cyprinus carpio (Common carp).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Cyprinus.  
 NCBI\_TaxId=7962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tsai P.-L., Chang G.-D., Huang C.-J.;  
 RT "Purification and cloning of carp fetuin."  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY225865; AA074862.1; -.  
 DR GO: GO:0005874; C:mitochondrion; IEA.  
 DR GO: GO:0004669; F:cysteine protease inhibitor activity; IEA.  
 DR GO: GO:0005198; F:structural molecule activity; IEA.  
 DR GO: GO:0007018; P:microtubule-based movement; IEA.  
 DR InterPro: IPR002453; Beta.tubulin.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; Cystatin; 1.  
 DR SMART: SM00043; CY; 1.  
 DR PROSITE: PS00228; TUBULIN B AUTOREG; 1.  
 SQ SEQUENCE 464 AA; 51698 MW; 7A54F71E4050895 CRC64;

Query Match 30.5%; Score 57; DB 13; Length 464;  
 Best Local Similarity 32.3%; Pred. No. 14;  
 Matches 10; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

Qy 2 DSLQWITDQYNKESDDKXHFRIFRVLKVGROV 32  
 Db 158 ESYKTAIQKFNKESDDKXHFRIFRVLKVGRISTQ 188

## RESULT 14

ID 080389 PRELIMINARY; PRT; 96 AA.  
 AC 080389;  
 DT 01-NOV-1998 (TEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)  
 DE Cystein protease inhibitor.  
 OS Cucumis sativus (Cucumber).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosidie I; Cucurbitales; Cucurbitaceae; Cucumis.  
 NCBI\_TaxId=3659;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamakawa S.;  
 RT "Cysteine proteinase inhibitor."  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB014760; BAA28867.1; -.  
 DR PIR: T07822; T07822.  
 DR GO: GO:0004669; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 DR SMART: SM00043; CY; 1.  
 DR PROSITE: PS00287; CYSTATIN; 1.  
 SQ SEQUENCE 96 AA; 10973 MW; 027252E14BB6C4F4 CRC64;

Query Match 29.9%; Score 56; DB 10; Length 96;  
 Best Local Similarity 33.3%; Pred. No. 3.2;

Matches 11; Conservative 6; Mismatches 14; Indels 2; Gaps 1;  
 Qy 1 KDSLQWITDQYNKESDDKXHFRIFRVLKVGROV 33  
 Db 21 KDIAEVAVAEYKNSQG--HHLTIVSILKCESQV 51

## RESULT 15

ID 09TY65 PRELIMINARY; PRT; 109 AA.  
 AC 09TY65;  
 DT 01-MAY-2000 (TEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)  
 DE Antigen maltose binding protein (Fragment).  
 OS Onchocerca volvulus.  
 OS Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Onchocerca.  
 NCBI\_TaxId=6282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94336252; PubMed=8058358;  
 RA Trenholme K.R., Tree T.I., Gillespie A.J., Guderian R., Maizels R.M.,  
 RA Bradley J.E.;  
 RT "Heterogeneity of IgG antibody responses to cloned Onchocerca volvulus  
 RT antigens in microfiladermia positive individuals from Bameraldas  
 RT Province, Ecuador."  
 RL Parasite Immunol. 16:201-209(1994).  
 DR EMBL: S71364; AAC60509.1; -.  
 DR GO: GO:0004669; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 DR SMART: SM00043; CY; 1.  
 DR PROSITE: PS00287; CYSTATIN; 1.  
 DR NON TER 1  
 FT SEQUENCE 109 AA; 12701 MW; 3C6B5E914D8082BB CRC64;

Query Match 29.9%; Score 56; DB 5; Length 109;  
 Best Local Similarity 45.5%; Pred. No. 3.7;  
 Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 12 NKESDDKXHFRIFRVLKVGROV 33  
 Db 24 NEQSNDEYHLMPIKLVSSQV 45

Search completed: March 18, 2004, 14:21:27  
 Job time : 21.9722 secs





```

DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
DE SC13delta.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF35481; AAL7192.1; -.
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
SQ
SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;

Query Match 92.0%; Score 172; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.5e-16;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDSLOWITTDQYNKESDDKHFRIFRVLKVRQ 32
Db 45 KDSLOWITTDQYNKESDDKHFRIFRVLKVRQ 76

RESULT 3
Q8K5A3 PRELIMINARY; PRT; 139 AA.
AC 08K5A3;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
DE Cystatin 11.
GN CST11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley;
RA Hamil K.G., Hall S.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF501290; AAM21709.1; -.
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART: SM00043; Cy; 1.
SQ
SEQUENCE 139 AA; 16686 MW; E1E36DB786B4D08C CRC64;

Query Match 62.0%; Score 116; DB 11; Length 139;
Best Local Similarity 48.6%; Pred. No. 1.7e-08;
Matches 17; Conservative 16; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KDSLOWITTDQYNKESDDKHFRIFRVLKVRQVTD 35
Db 45 KETLEYTEBYNKSEDLNFRILRIILKIKRKMTN 79

RESULT 4
Q9DAP1 PRELIMINARY; PRT; 141 AA.
AC Q9DAP1;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
DE 1700006C19R1k protein.
GN 1700006C19R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Mazzarelli J., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseuki S.,
RA Hayaehizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK005665; BAB24175.1; -.
DR HSSP: P01038; ICEW.
DR MGD; MGI:1916544; 1700006C19R1k.
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART: SM00043; Cy; 1.
SQ
SEQUENCE 141 AA; 16811 MW; C20FA0DB1AC378C CRC64;

Query Match 44.4%; Score 83; DB 11; Length 141;
Best Local Similarity 45.5%; Pred. No. 0.00077;
Matches 15; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 3 SLOWITTDQYNKESDDKHFRIFRVLKVRQVTD 35
Db 47 ALWYAMKEYNKASNDLNFRLVILKSEQEITD 79

RESULT 5
Q80ZNS PRELIMINARY; PRT; 141 AA.
AC Q80ZNS;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE RIKEN cDNA 1700006C19 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Testicle;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC048681; AAH48681.1; -.
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR InterPro: IPR003243; Cystatin_C/M.
DR Pfam: PF00031; cystatin; 1.
DR ProDom: PD001231; Cystatin_C/M; 1.
DR SMART: SM00043; Cy; 1.
SQ
SEQUENCE 141 AA; 16825 MW; C20FA0DBA884951F CRC64;

Query Match 44.4%; Score 83; DB 11; Length 141;
Best Local Similarity 45.5%; Pred. No. 0.00077;
Matches 15; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 3 SLOWITTDQYNKESDDKHFRIFRVLKVRQVTD 35

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 20.9722 Seconds  
(without alignments)  
526.560 Million cell updates/sec

Title: US-09-941-314-8  
Perfect score: 187  
Sequence: 1 KDSIQWITDQYNKESDDKHFRIFRVLKVGQVTD 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mhc:\*  
8: SP\_Organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP\_vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_virus:\*  
16: SP\_bacteriap:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187	100.0	138	4 Q8WXU6	Q8WXU6 homo sapien
2	172	92.0	103	4 Q8WXU5	Q8WXU5 homo sapien
3	116	62.0	139	11 Q8K5A3	Q8K5A3 rattus norv
4	83	44.4	141	11 Q9DAP1	Q9DAP1 mus musculu
5	83	44.4	141	11 Q80ZM5	Q80ZM5 mus musculu
6	77	41.2	112	13 Q98SR4	Q98SR4 acipenser s
7	77	41.2	112	13 Q98SR3	Q98SR3 acipenser s
8	69	36.9	81	6 Q29212	Q29212 sus scrofa
9	67	35.8	140	11 Q9EPX9	Q9EPX9 mus musculu
10	61	32.6	148	5 Q9NH95	Q9NH95 litomosoid
11	60.5	32.4	4530	5 Q812K6	Q812K6 plasmodium
12	57	30.5	161	5 Q16159	Q16159 brugia mala
13	57	30.5	464	13 Q801Z5	Q801Z5 cyprinus ca
14	56	29.9	96	10 Q80389	Q80389 cucumis sat
15	56	29.9	109	5 Q9TY65	Q9TY65 onchocerca
16	56	29.9	222	12 Q39107	Q39107 bunyavirus

17	54	28.9	222	12 Q39106	Q39106 bunyavirus
18	54	28.9	235	12 Q8JPR0	Q8JPR0 bunyavirus
19	54	28.9	400	16 Q9ZDT8	Q9ZDT8 rickettsia
20	54	28.9	518	10 Q22157	Q22157 arabidopsis
21	54	28.9	878	10 Q8RXD6	Q8RXD6 arabidopsis
22	54	28.9	1320	5 Q81UJ3	Q81UJ3 plasmodium
23	54	28.9	2646	12 Q9WPZ7	Q9WPZ7 gill-associ
24	53	28.3	132	2 Q86214	Q86214 exiguobacte
25	53	28.3	194	5 Q61819	Q61819 caenorhabdi
26	53	28.3	205	5 Q817Y7	Q817Y7 caenorhabdi
27	53	28.3	319	16 Q9CJ05	Q9CJ05 lactococcus
28	53	28.3	505	5 Q81U00	Q81U00 caenorhabdi
29	53	28.3	551	11 Q61192	Q61192 mus musculu
30	53	28.3	551	11 Q8VIG7	Q8VIG7 mus musculu
31	53	28.3	564	11 Q8C6B2	Q8C6B2 mus musculu
32	53	28.3	651	2 Q9RAI8	Q9RAI8 mycobacter
33	53	28.3	726	2 Q9IAE4	Q9IAE4 azospirillum
34	53	28.3	844	5 Q86S89	Q86S89 polypilaetro
35	53	28.3	3429	11 Q08614	Q08614 mus musculu
36	52	27.8	109	16 Q7TU55	Q7TU55 prochloroco
37	52	27.8	132	2 Q9F4C0	Q9F4C0 bacillus ma
38	52	27.8	132	2 Q70107	Q70107 bacillus ce
39	52	27.8	213	10 Q8LK68	Q8LK68 cucumis sat
40	52	27.8	272	5 Q9XW68	Q9XW68 caenorhabdi
41	52	27.8	279	16 Q8R5P1	Q8R5P1 fusobacteri
42	52	27.8	428	10 Q9FYQ9	Q9FYQ9 arabidopsis
43	52	27.8	450	16 Q92IU9	Q92IU9 rickettsia
44	52	27.8	1200	11 Q921B9	Q921B9 mus musculu
45	51.5	27.5	133	11 Q9D264	Q9D264 mus musculu

## ALIGNMENTS

## RESULT 1

Q8WXU6 PRELIMINARY; PRT; 138 AA.  
ID Q8WXU6  
AC Q8WXU6;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE SCT3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;  
RT "SCT3: A novel epididymal specific member of the cystatin family.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF35480; F: cysteine protease inhibitor activity; IEA.  
DR GO: GO:0004869; F: cysteine protease inhibitor activity; IEA.  
DR InterPro: IPR000010; Cysteatin.  
DR Pfam: PF00031; Cysteatin; 1.  
DR SMART: SM00043; CY; 1.  
SQ SEQUENCE 138 AA; 16506 MW; E49440ACA3585C64 CRC64;

Query Match 100.0%; Score 187; DB 4; Length 138;  
Best Local Similarity 100.0%; Pred. No. 1.6e-18;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSIQWITDQYNKESDDKHFRIFRVLKVGQVTD 35

DB 45 KDSIQWITDQYNKESDDKHFRIFRVLKVGQVTD 79

## RESULT 2

ID Q8WXU5 PRELIMINARY; PRT; 103 AA.  
AC Q8WXU5;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)



KW Thiol protease inhibitor; Amyloid; Signal; Disease mutation;  
KW Polymorphism; 3D-structure.  
FT SIGNAL 1 26  
FT CHAIN 27 146 CYSTATIN C.  
FT ACT SITE 37 37 REACTIVE SITE.  
FT SITE 81 85 SECONDARY AREA OF CONTACT.  
FT DISULFID 99 109  
FT DISULFID 123 143

Query Match 36.9%; Score 69; DB 1; Length 146;  
Best Local Similarity 35.5%; Pred. No. 0.016;  
Matches 11; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 3 SLOWITDQVKNESDDKYHFRIFVLKVRQV 33  
DB 52 ALDFAVGEYKNSNDMTHSRALQVVRKQI 82

RESULT 13  
CYTC RAT  
ID CYTC RAT STANDARD; PRT; 127 AA.  
AC P14841;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin C precursor (Fragment).  
GN C573.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Buffalo;  
RX MEDLINE=90092122; PubMed=2689174;  
RA Cole T., Dickson P.W., Enard F., Averill F., Risbridger G.,  
RA Gauthier F., Schreiber G.;  
RT "The cDNA structure and expression analysis of the genes for the  
RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin  
RT in rat brain.";  
RL Eur. J. Biochem. 186:35-42 (1989).  
RN [2]  
RP SEQUENCE OF 8-127.  
RX MEDLINE=90380276; PubMed=2400577;  
RA Enard F., Enard A., Faucher D., Capony J.-P., Derancourt J.,  
RA Brillard M., Gauthier F.;  
RT "Rat cystatin C: the complete amino acid sequence reveals a site for  
RT N-glycosylation.";  
RL Biol. Chem. Hoppe-Seyler 371:161-166 (1990).  
RN [3]  
RP SEQUENCE OF 8-49.  
RX MEDLINE=88313020; PubMed=3044831;  
RA Enard A., Enard F., Faucher D., Gauthier F.;  
RT "Two rat homologues of human cystatin C.";  
RL FEBS Lett. 236:475-478 (1988).  
RN [4]  
RP SEQUENCE OF 8-20.  
RC TISSUE=Sertoli cells;  
RX MEDLINE=9225121; PubMed=1563513;  
RA Enard A., Enard F., Guillou F., Gauthier F.;  
RT "Production of the cysteine proteinase inhibitor cystatin C by rat  
RT Sertoli cells.";  
RL FEBS Lett. 300:131-135 (1992).  
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
CC thought to serve an important physiological role as a local  
CC regulator of this enzyme activity. Known to inhibit cathepsin B,  
CC H, and L.  
CC -1- SIMILARITY: Belongs to the cystatin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X16957; CAA34831.1; -.  
CC DR PIR; S07085; S07085.  
CC DR PIR; S10587; S10587.  
CC DR HSP; P01034; I96.  
CC DR InterPro; IPR000010; Cystatin.  
CC DR Pfam; PF00031; cystatin; 1.  
CC DR SMART; SM00043; CY; 1.  
CC DR PROSITE; PS00287; CYSTATIN; 1.  
CC KW Thiol protease inhibitor; Signal.  
CC FT NON TER 1 1  
CC FT SIGNAL <1 7  
CC FT CHAIN 8 127 CYSTATIN C.  
CC FT ACT SITE 18 18 REACTIVE SITE.  
CC FT SITE 62 66 SECONDARY AREA OF CONTACT.  
CC FT DISULFID 80 90 BY SIMILARITY.  
CC FT DISULFID 104 124 BY SIMILARITY.  
CC FT CONFLICT 25 25 A -> E (IN REF. 2).  
CC SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;  
Query Match 35.8%; Score 67; DB 1; Length 127;  
Best Local Similarity 35.5%; Pred. No. 0.026;  
Matches 11; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 3 SLOWITDQVKNESDDKYHFRIFVLKVRQV 33  
DB 33 ALDFAVSEYKNSNDAYHSGRAIQVVRKQL 63

RESULT 14  
CYT CYPCA  
ID CYT CYPCA STANDARD; PRT; 129 AA.  
AC P35481;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin precursor (Ovarian cystatin) (P12).  
OS Cyprinus carpio (Common carp).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Cyprinus.  
OX NCBI\_TaxID=7962;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-57; 74-80 AND 92-129.  
RC TISSUE=Ovary;  
RX MEDLINE=96208938; PubMed=8829807;  
RA Tsai Y.-J., Chang G.-D., Huang C.-J., Chang Y.-S., Huang F.-L.;  
RT "Purification and molecular cloning of carp ovarian cystatin.";  
RL Comp. Biochem. Physiol. 113B:573-580 (1996).  
CC -1- FUNCTION: Cysteine proteinase inhibitor.  
CC -1- TISSUE SPECIFICITY: Detected in ovary but not in all the other  
CC examined tissues.  
CC -1- PTM: THIS PROTEIN IS FURTHER PROCESSED BY AN INTRAMOLECULAR  
CC CLEAVAGE TO PRODUCE TWO CHAINS LINKED BY A DISULFIDE BRIDGE.  
CC -1- SIMILARITY: Belongs to the cystatin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L23572; AAB48011.1; -.  
CC HSP; P01038; ICEW.  
CC DR InterPro; IPR000010; Cystatin.  
CC DR Pfam; PF00031; cystatin; 1.  
CC DR SMART; SM00043; CY; 1.  
CC DR PROSITE; PS00287; CYSTATIN; 1.



[4] SEQUENCE FROM N.A.  
RP  
RX MEDLINE-89350949; PubMed-2764935;  
RA Saïton E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,  
RA Isemura S., Sanada K.  
RT "The human cystatin C gene (CSTR) is a member of the cystatin gene  
RT family which is localized on chromosome 20."  
RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).  
RN [5]



RL Biochem. J. 217:813-817(1984).  
RN [3]  
RP SEQUENCE OF 24-139.  
RX MEDLINE=84110059; PubMed=6662498;  
RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,  
MAchleidt W.;  
RA "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
of cystatin from chicken egg white";  
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).  
RN [4]  
RP CHARACTERIZATION OF PROTEIN.  
RX MEDLINE=83256421; PubMed=6409085;  
RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,  
Sunter D.C., Barrett A.J.;  
RA "Cystatin, a protein inhibitor of cysteine proteinases. Improved  
purification from egg white, characterization, and detection in  
chicken serum";  
RL Biochem. J. 211:129-138(1983).  
RN [5]  
RP DISULFIDE BONDS  
RA Grubb A., Loeffberg H., Barrett A.J.;  
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
cystatin";  
RL FEBS Lett. 170:370-374(1984).  
RN [6]  
RP PHOSPHORYLATION  
RX MEDLINE=89252033; PubMed=2721673;  
RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,  
Bode W.;  
RT "The cysteine proteinase inhibitor chicken cystatin is a  
phosphoprotein";  
RL FEBS Lett. 248:162-168(1989).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=89052676; PubMed=3191914;  
RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshikov A.,  
Brzin J., Kos J., Turk V.;  
RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and  
its possible mode of interaction with cysteine proteinases";  
RL EMBO J. 7:2593-2599(1988).  
RN [8]  
RP STRUCTURE BY NMR.  
RX MEDLINE=94087719; PubMed=8263912;  
RA Dieckmann T., Mitschang L., Hofmann M., Kos J., Turk V.,  
Auerwald E.A., Jeanicke R., Oschkinat H.;  
RT "The structures of native phosphorylated chicken cystatin and of a  
recombinant unphosphorylated variant in solution";  
RL J. Mol. Biol. 234:1048-1059(1993).  
CC -I- FUNCTION: This protein binds tightly to and inhibits a variety of  
thiol proteases including ficin, papain, and cathepsins B, C, H,  
and L. Although isolated from egg white, it is also present in  
serum.  
CC -I- SIMILARITY: Belongs to the cystatin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J05077; AAA48744.1; -;  
DR PIR; A34456; UDCH.  
DR PDB; 1CEW; 31-JAN-94.  
DR PDB; 1A67; 27-MAY-98.  
DR PDB; 1A90; 17-JUN-98.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00031; cystatin; 1.  
DR SMART; SM00043; CY; 1.  
DR PROSITE; PS00287; CYSTATIN; 1.  
KW Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.  
FT SIGNAL 1 23

FT CHAIN 24 139 CYSTATIN.  
FT ACT SITE 32 32 REACTIVE SITE.  
FT SITE 76 80 SECONDARY AREA OF CONTACT.  
FT DISULFID 94 104  
FT DISULFID 118 138  
FT MOD RES 103 103 PHOSPHORYLATION (PARTIAL).  
FT STRAND 35 36  
FT TURN 39 40  
FT HELIX 42 51  
FT TURN 52 52  
FT HELIX 53 56  
FT TURN 57 58  
FT STRAND 63 77  
FT STRAND 81 95  
FT TURN 96 97  
FT TURN 99 100  
FT HELIX 101 108  
FT STRAND 115 125  
FT TURN 126 129  
FT STRAND 130 139  
SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;  
Query Match 39.0%; Score 73; DB 1; Length 139;  
Best Local Similarity 41.9%; Pred. No. 0.0044;  
Matches 13; Conservative 9; Mismatches 9; Indels 0; Gaps 0;  
QY 3 SLOWTDQYNKSDDKYHFRVLKVRQV 33  
DB 47 ALQFAMAEYNRASNDKYSRVVRVISAKRQL 77  
RESULT 7  
CYTC\_RABIT STANDARD; PRT; 148 AA.  
AC Q97862;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin C precursor.  
GN CST3.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Japanese white; TISSUE=Bone;  
RX MEDLINE=98424349; PubMed=9753427;  
RA Kobori M., Ikeda Y., Nara H., Kato M., Kumegawa M., Nojima H.,  
Kawashima H.;  
RT "Large scale isolation of osteoclast-specific genes by an improved  
method involving the preparation of a subtracted cDNA library";  
RL Genes Cells 3:459-475(1998).  
CC -I- FUNCTION: This is a thiol proteinase inhibitor.  
CC -I- SIMILARITY: Belongs to the cystatin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB009342; BAA75921.1; -;  
DR HSSP; P01034; 1G96.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00031; cystatin; 1.  
DR SMART; SM00043; CY; 1.  
DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
KW Thiol protease inhibitor; Signal.  
FT SIGNAL 1 28 POTENTIAL.  
FT CHAIN 29 148 CYSTATIN C.

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin SA precursor (Cystatin S5).  
 GN CST2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88185836; PubMed=3446578;  
 RA Saitoh E., Kim H.-S., Smithies O., Maeda N.;  
 RT "Human cysteine-proteinase inhibitors: nucleotide sequence analysis  
 of three members of the cystatin gene family.";  
 RL Gene 61:329-338(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehaeslahti M.H., Leverhulme M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prachinagam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycomore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RN Nature 414:865-871(2001).  
 RN [3]  
 RP SEQUENCE OF 21-40.  
 RC TISSUE=Saliva;  
 RX MEDLINE=92138674; PubMed=1778989;  
 RA Isemura S., Saitoh E., Sanada K., Minakata K.;  
 RT "Identification of full-sized forms of salivary (S-type) cystatins  
 (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of  
 cystatin S) in human whole saliva and determination of phosphorylation  
 sites of cystatin S.";  
 RL J. Biochem. 110:648-654(1991).  
 RN [4]  
 RP SEQUENCE OF 25-141.  
 RX MEDLINE=88139220; PubMed=3436950;  
 RA Isemura S., Saitoh E., Sanada K.;  
 RT "Characterization and amino acid sequence of a new acidic cysteine  
 proteinase inhibitor (cystatin SA) structurally closely related to  
 cystatin S, from human whole saliva.";  
 RL J. Biochem. 102:693-704(1987).  
 RN [5]  
 RP PRELIMINARY SEQUENCE OF 25-141.  
 RA Isemura S., Saitoh E., Sanada K., Isemura M., Ito S.;  
 RT "Characterization and amino acid sequence of a new acidic cysteine  
 proteinase inhibitor (cystatin SA) structurally closely related to  
 cystatin S, from human whole saliva.";  
 RL (in) Turk V. (eds.);  
 RL Cysteine proteinases and their inhibitors, pp.497-505,  
 RL Walter de Gruyter, Berlin and New York (1986).  
 RN [6]  
 RP SEQUENCE OF 25-141 FROM N.A.

RX MEDLINE=89076505; PubMed=3202964;  
 RA Saitoh E., Isemura S., Sanada K., Kim H.-S., Smithies O., Maeda N.;  
 RT "Cystatin superfamily. Evidence that family II cystatin genes are  
 evolutionarily related to family III cystatin genes.";  
 RL Biol. Chem. Hoppe-Seyler 369:191-197(1988).  
 CC -!- FUNCTION: Thiol protease inhibitor.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M19673; AAA361116.1; --  
 CC EMBL; M19671; AAA361116.1; JOINED.  
 CC EMBL; M19672; AAA361116.1; JOINED.  
 CC EMBL; AL591074; CAC94784.1; --  
 CC PIR; B29632; B29632.  
 CC HSSP; P01034; IG96.  
 CC Genew; HGNC:2474; CST2.  
 CC MIM; 123856; --  
 CC GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.  
 CC InterPro; IPR000010; Cystatin.  
 CC Pfam; PF00031; cystatin; 1.  
 CC SMART; SM00043; CY; 1.  
 CC PROSITE; PS00287; CYSTATIN; 1.  
 KW Thiol protease inhibitor; Signal; Multigene family.  
 FT SIGNAL 1 20  
 FT CHAIN 21 141 CYSTATIN SA.  
 FT ACT SITE 32 32 REACTIVE SITE.  
 FT SITE 76 80 SECONDARY AREA OF CONTACT.  
 FT DISULFID 94 104 BY SIMILARITY.  
 FT DISULFID 118 138 BY SIMILARITY.  
 SQ SEQUENCE 141 AA; 16445 MW; B54915B1B977AA2 CRC64;  
 Query Match 39.6%; Score 74; DB 1; Length 141;  
 Best Local Similarity 35.5%; Pred. No. 0.0032;  
 Matches 11; Conservative 12; Mismatches 8; Indels 0; Gaps 0;  
 QY 3 SLOWITDQYNKESDDKHFRIPLVLRVLRQVQ 33  
 Db 47 ALHFVISEYNKATEDEYRRLRLVLRAREQI 77  
 RESULT 6  
 CYT CHICK  
 ID CYT CHICK STANDARD; PRT; 139 AA.  
 AC P01038;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin precursor (Egg-white cystatin).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90008873; PubMed=2793849;  
 RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;  
 RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,  
 and tissue distribution.";  
 RL J. Biol. Chem. 264:17164-17169(1989).  
 RN [2]  
 RP SEQUENCE OF 24-139  
 RX MEDLINE=84178305; PubMed=6712597;  
 RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;  
 RT "Cystatin. Amino acid sequence and possible secondary structure.";

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.W., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 CC -!- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epithelium.  
 CC -!- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF059244; AAC14707.1; -;  
 CC EMBL; AL109954; CAB64234.1; -;  
 CC HSSP; P01034; 1G96.  
 CC Genew; HGNC:2480; CST8.  
 CC GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.  
 CC InterPro; IPR000010; Cystatin.  
 CC Pfam; PF00031; cystatin; 1.  
 CC SMART; SM00043; Cy; 1.  
 CC Thiol protease inhibitor; Signal; Polymorphism.  
 FT SIGNAL 1 21  
 FT CHAIN 22 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 FT PROTEIN.  
 FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 95 105 BY SIMILARITY.  
 FT DISULFID 119 139 BY SIMILARITY.  
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 142 142 A -> P (in dbSNP:1054633).  
 FT FTId=VAR 014527.  
 SQ SEQUENCE 142 AA; 16275 MW; 9A351275780F4ECD CRC64;  
 Query Match 43.9%; Score 82; DB 1; Length 142;  
 Best Local Similarity 45.7%; Pred. No. 0.00026;  
 Matches 16; Conservative 7; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLVKQVQRTD 35  
 DB 46 KCLWFAMQEQYNKESDKYVFLVKTLQALQVNT 80  
 RESULT 4  
 CST8\_MOUSE STANDARD; PRT; 142 AA.  
 AC P32766; O89102;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-  
 DE related epididymal specific protein) (Cystatin 8).  
 GN CST8 OR CR5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;

RN SEQUENCE FROM N.A.  
 RP STRAIN=C3H, and CD-1;  
 RX MEDLINE=99247899; PubMed=10229662;  
 RA Cornwall G.A., Hsia N., Sutton H.G.;  
 RT "Structure, alternative splicing and chromosomal localization of the  
 RT cystatin-related epididymal spermatogenic gene.";  
 RL Biochem. J. 340:85-93(1999).  
 RN [2]  
 RP SEQUENCE OF 4-142 FROM N.A.  
 RC TISSUE=Episididymis;  
 RX MEDLINE=93078799; PubMed=1280328;  
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;  
 RT "The CR5 gene: a unique testis-regulated gene related to the cystatin  
 RT family is highly restricted in its expression to the proximal region  
 RT of the mouse epididymis.";  
 RL Mol. Endocrinol. 6:1653-1664(1992).  
 CC -!- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epithelium.  
 CC -!- INDUCTION: Testicular factors or hormones other than androgens  
 CC present in the testicular fluid may be involved in the regulation  
 CC of CR5 gene expression.  
 CC -!- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF091503; AAC61754.1; -;  
 CC EMBL; AF090691; AAC36316.1; -;  
 CC EMBL; S49926; AAC35390.1; -;  
 CC PIR; A45361; A45361.  
 CC HSSP; P01034; 1G96.  
 CC MGD; MGI:107161; Cst8.  
 CC InterPro; IPR000010; Cystatin.  
 CC Pfam; PF00031; cystatin; 1.  
 CC SMART; SM00043; Cy; 1.  
 CC Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 FT PROTEIN.  
 FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 95 105 BY SIMILARITY.  
 FT DISULFID 119 139 BY SIMILARITY.  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 4 15 PMLSLILFLIIP -> GFDEQVGSQK (IN REF. 2).  
 SQ SEQUENCE 142 AA; 16288 MW; 50B446E98F6673E CRC64;  
 Query Match 41.2%; Score 77; DB 1; Length 142;  
 Best Local Similarity 40.0%; Pred. No. 0.0013;  
 Matches 14; Conservative 9; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLVKQVQRTD 35  
 DB 46 KQCVWFAMQEQYNKESDKYVFLVKTLQALQVNT 80  
 RESULT 5  
 CYTT\_HUMAN STANDARD; PRT; 141 AA.  
 ID CYTT\_HUMAN  
 AC P09228; Q9UCQ7;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AL096677; CAC13170.1; -;  
 DR EMBL; AL096677; CAC17423.1; -;  
 DR HSSP; P01038; 1A90.  
 DR Genew; HGNC:15959; CST11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
 KW Thiol protease inhibitor; Signal; Alternative splicing.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 137 CYSTATIN 11.  
 FT SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 93 101 BY SIMILARITY.  
 FT DISULFID 114 134 BY SIMILARITY.  
 FT CARBOHYD 131 131 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT VARSPLIC 76 110 Missing (in isoform 2).  
 FT FTId=VSP 001260.  
 SQ SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 187; DB 1; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 9.3e-19;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSLOWITQYKESDDKYHFRIFRVLKVRQVTD 35  
 DB 44 KDSLOWITQYKESDDKYHFRIFRVLKVRQVTD 78

## RESULT 2

CS11\_MOUSE STANDARD; PRT; 139 AA.  
 AC Q9D269;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin 11 precursor.  
 GN CS11.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Hayashaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.  
 RT Functional annotation of a full-length mouse cDNA collection.;  
 RL Nature 409:685-690(2001).  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AK020300; BAB32061.1; -;  
 DR HSSP; P01034; 1G96.  
 DR MGD; MGI:1925490; Cst11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
 KW Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT CHAIN 29 139 CYSTATIN 11.  
 FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 94 102 BY SIMILARITY.  
 FT DISULFID 115 135 BY SIMILARITY.  
 FT CARBOHYD 134 134 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 139 AA; 16217 MW; F228D9815FA32640 CRC64;

Query Match 64.2%; Score 120; DB 1; Length 139;  
 Best Local Similarity 55.9%; Pred. No. 1.5e-09;  
 Matches 19; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 KDSLOWITQYKESDDKYHFRIFRVLKVRQVTD 34  
 DB 45 KETLEYVTDYKNSDLYNFRILRLIKMKQVT 78

## RESULT 3

CS18\_HUMAN STANDARD; PRT; 142 AA.  
 AC O60676;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin  
 DE 8).  
 GN CST8 OR CRSS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=95344753; PubMed=7619504;  
 RA Cornwall G.A., Hann S.R.;  
 RT "Transient appearance of CRSS protein during spermatogenesis and  
 RT caput epididymal sperm maturation.";  
 RL Mol. Reprod. Dev. 41:37-46(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Grahame D.V., Griffiths C., Griffiths J.A., Holden J.L., Howden P.J.,  
 RA Hammond S., Hunt J.L., Heath P.D., Ho S., Holden J.L., Johnson P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehesvaara M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 4.16667 Seconds  
(without alignments)  
437.389 Million cell updates/sec

Title: US-09-941-314-8

Perfect score: 187

Sequence: 1 KDSLQWITQYNKESDDKYHFRIFRVLKQVQVTD 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	187	100.0	137	CS11 HUMAN	Q9h12 homo sapien
2	120	64.2	139	CS11 MOUSE	Q9d269 mus musculus
3	82	43.9	142	CS18 HUMAN	O60676 homo sapien
4	77	41.2	142	CS18 MOUSE	P12766 mus musculus
5	74	39.6	141	CYT HUMAN	P09228 homo sapien
6	73	39.0	139	CYT CHICK	F01038 gallus galli
7	72	38.5	148	CYT RABIT	O97862 oryctolagus
8	71	38.0	116	CYT COTJA	P81061 coturnix co
9	70	37.4	146	CYT MACMU	O19092 macaca mula
10	70	37.4	146	CYT SALSC	O19093 sainiri sci
11	70	37.4	148	CYT BOVIN	P01035 bos taurus
12	69	36.9	146	CYT HUMAN	P01034 homo sapien
13	67	35.8	127	CYT RAT	P14841 rattus norv
14	67	35.8	129	CYT CYPCA	P35481 cyprinus ca
15	67	35.8	140	CYT MOUSE	P21460 mus musculus
16	64	34.2	142	CS18 RAT	O88969 rattus norv
17	61	32.6	141	CYT RAT	P19313 rattus norv
18	60	32.1	130	CYT ONCKE	Q98967 oncorhynch
19	60	32.1	130	CYT ONCMY	Q91195 oncorhynch
20	58	31.0	141	CYT HUMAN	P01037 homo sapien
21	58	31.0	141	CYT HUMAN	P01036 homo sapien
22	56	29.9	162	CYT ONCVO	P22085 onchocerca
23	56	29.9	165	CS18 HUMAN	Q9h14 homo sapien
24	54	28.9	235	NCAP BUNLC	P04873 bunyavirus
25	53	28.3	111	CYT BITAR	P08935 bitis ariet
26	53	28.3	132	MERR BACCE	P22853 bacillus ce
27	53	28.3	474	SYE VIBPA	Q87rl6 vibrio para
28	52.5	28.1	255	DA82 YEAST	P21705 saccharomyc
29	52.5	28.1	735	DHDM HTPSX	Q48303 hyphomicrob
30	52	27.8	504	GUNW ERWCA	O59395 erwinia car
31	52	27.8	3433	UTRO HUMAN	P46939 homo sapien
32	51	27.3	474	SYE VIBVY	Q7mmw8 vibrio vuln
33	51	27.3	505	VP5_AHSV4	Q02168 african hor

RESULT 1  
CS11\_HUMAN STANDARD; PRT; 137 AA.

AC Q9H12; Q9H113.

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cystatin 11 precursor.

GN CST11 OR CST8L

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,

RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,

RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,

RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,

RA Lehaeslao M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConachie L.J., McMay K., McMurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,

RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,

RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,

RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,

RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,

RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,

RA Wilming L., Wray F.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.";

RL Nature 414:865-871(2001).

CC -1- SUBCELLULAR LOCATION: Secreted (Potential).

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1;

CC IsoId=Q9H112-1; Sequence=Displayed;

CC Name=2;

CC IsoId=Q9H112-2; Sequence=VSP\_001260;

CC Note=No experimental confirmation available;

CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

A;Status: not compared with conceptual translation  
A:Molecule type: DNA  
A;Residues: 21-141 <SA2>  
R;Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.  
Gene 61, 329-338, 1987  
A;Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three mem  
A;Reference number: A91589; MUID:88185836; PMID:3445578  
A;Accession: A29632  
A:Molecule type: DNA  
A;Residues: 1-86, 'I', 88-141 <SA1>  
R;Isemura, S.; Saitoh, E.; Sanada, K.  
FEBS Lett. 198, 145-149, 1986  
A;Title: Characterization of a new cysteine proteinase inhibitor of human saliva, cystat  
A;Reference number: A01273; MUID:86164938; PMID:3514272  
A;Accession: A01273  
A:Molecule type: protein  
A;Residues: 29-141 <ISE>  
R;Ramasubbun, N.; Reddy, M.S.; Bergey, E.J.; Harasathy, G.G.; Soni, S.D.; Levine, M.J.  
Biochem. J. 280, 341-352, 1991  
A;Title: Large-scale purification and characterization of the major phosphoproteins and  
A;Reference number: S19279; MUID:92082469; PMID:1747107  
A;Accession: S19279  
A;Status: preliminary  
A:Molecule type: protein  
A;Residues: 21-55 <RAM>  
C;Comment: Human saliva appears to contain several cysteine proteinase inhibitors that a  
ences. Cystatin SN, with a pI of 7.5, is a much better inhibitor of papain and dipeptyl  
C;Genetics:  
A;Gene: GDB:CST1  
A;Cross-references: GDB:119815; OMIM:123855  
A;Map position: 20p11.2-20p11.2  
C;Superfamily: cystatin; cystatin homology  
C;Keywords: cysteine proteinase inhibitor; extracellular protein; saliva  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-141/Product: cystatin SA-I #status experimental <MAT1>  
F;23-141/Product: cystatin SN #status experimental <MAT2>  
F;30-141/Domain: cystatin homology <CYS>  
F;76-80/Region: inhibitory #status predicted  
F;94-104,118-138/Disulfide bonds: #status predicted

Query Match 31.0%; Score 58; DB 1; Length 141;  
Best Local Similarity 36.7%; Pred. No. 1.5;  
Matches 11; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRIFRVLKVQRQ 32  
: : : : :  
DB 47 ALHPAISEYNKATKDYYRRPLRLVLRARQQ 76  
: : : : :  
Search completed: March 18, 2004, 14:23:14  
Job time : 6.73611 secs

A:Molecule type: protein  
A:Residues: 29-38;76-80 <BIO>  
A>Note: forms RSC-1, RSC-2 and RSC-3 with differing amino-terminals were found; a form B  
C:Genetics:  
A:Gene: Cyss  
A:Introns: 76/3; 114/3  
C:Superfamily: cystatin; cystatin homology  
C:Keywords: cysteine proteinase inhibitor  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:28-141/Product: cystatin S #status experimental <MAT>  
F:29-141/Product: cystatin S, form RSC-3 #status predicted <MAT>  
F:30-141/Domain: cystatin homology <CYS>  
F:31-141/Product: cystatin S, form RSC-2 #status predicted <MAT>  
F:32-141/Product: cystatin S, form RSC-1 #status predicted <MAT>  
F:94-104,118-138/Disulfide bonds: #status experimental

Query Match 32.4%; Score 61; DB 2; Length 141;  
Best Local Similarity 34.4%; Pred. No. 0.58;  
Matches 11; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 2 DSLQWITQYNKESDDKYHFRIFRVLKQVQV 33  
DB 46 EALYAVNEXNKSDDLYLSRVVVEKDVQKQV 77

RESULT 11  
JC2040  
cystatin - chum salmon  
N:Alternate names: cysteine proteinase inhibitor  
C:Species: Oncorhynchus keta (Chum salmon)  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 16-Jul-1999  
C:Accession: JC2040  
R:Koida, Y.; Noso, T.  
BioSci. Biotechnol. Biochem. 58, 164-169, 1994  
A:Title: The complete amino acid sequence of pituitary cystatin from chum salmon.  
A:Reference number: JC2040; MUID:94162738; PMID:7764512  
A:Accession: JC2040  
A:Molecule type: protein  
A:Residues: 1-111 <ROI>  
C:Comment: The intracellular role of this protein is the inhibition of intralysosomal ph  
C:Superfamily: cystatin; cystatin homology  
C:Keywords: cysteine proteinase inhibitor  
F:2-111/Domain: cystatin homology <CYS>  
F:48-52/Region: inhibitory  
F:89-109/Disulfide bonds: #status experimental

Query Match 32.1%; Score 60; DB 1; Length 111;  
Best Local Similarity 30.3%; Pred. No. 0.61;  
Matches 10; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 KDSLQWITQYNKESDDKYHFRIFRVLKQVQV 33  
DB 17 RDLQPAVVEHNKKTNDMFVRQVAVVNAQKQV 49

RESULT 12  
JC4918  
cystatin precursor - chum salmon  
C:Species: Oncorhynchus keta (Chum salmon)  
C:Date: 26-Sep-1996 #sequence\_revision 01-Nov-1996 #text\_change 21-Jan-2000  
C:Accession: JC4918  
R:Yamashita, M.; Konagaya, S.  
J. Biochem. 120, 483-487, 1996  
A:Title: Molecular cloning and gene expression of chum salmon cystatin.  
A:Reference number: JC4918; MUID:97058289; PMID:8902609  
A:Accession: JC4918  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-132 <YAM>  
A:Cross-references: DDBJ:DB6628  
A:Experimental source: liver  
C:Comment: This protein is a homolog of mammalian cystatin C, and acts as an extracellular  
C:Superfamily: cystatin; cystatin homology

C:Keywords: liver  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:23-132/Domain: Cystatin homology <CYS>

Query Match 32.1%; Score 60; DB 2; Length 132;  
Best Local Similarity 30.3%; Pred. No. 0.74;  
Matches 10; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 KDSLQWITQYNKESDDKYHFRIFRVLKQVQV 33  
DB 38 RDLQPAVVEHNKKTNDMFVRQVAVVNAQKQV 70

RESULT 13  
E64005  
hypothetical protein HI0284 - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 26-Aug-1999  
C:Accession: E64005  
R:Rieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: E64005  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-139 <TIGR>  
A:Cross-references: GB:U32714; GB:L42023; NID:g1573241; PID:g1573253; TIGR:HI0284  
C:Genetics:  
A:Start codon: GTG  
C:Superfamily: Haemophilus influenzae hypothetical protein HI0284

Query Match 31.6%; Score 59; DB 2; Length 139;  
Best Local Similarity 45.0%; Pred. No. 1.1;  
Matches 9; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 15 SDDKYHFRIFRVLKQVQV 34  
DB 91 NDKKFFKIFILKIQRKLT 110

RESULT 14  
UDHUP1  
cystatin S precursor - human  
N:Alternate names: cystatin SA-III; salivary acidic protein-1  
C:Species: Homo sapiens (man)  
C:Date: 25-Feb-1985 #sequence\_revision 08-Feb-1996 #text\_change 16-Jul-1999  
C:Accession: S17667; S16500; A01272; A29603; S19280; A56608  
R:Bobek, L.A.; Aguirre, A.; Levine, M.J.  
Biochem. J. 278, 627-635, 1991  
A:Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ an  
A:Reference number: S17667; MUID:91378918; PMID:1898352  
A:Accession: S17667  
A:Molecule type: mRNA  
A:Residues: 1-141 <BOB>  
A:Cross-references: EMBL:X54667; NID:g30365; PIDN:CAA38478.1; PID:g30366  
R:Larkin, M.S.; Jensen, J.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.  
Arch. Biochem. Biophys. 288, 664-670, 1991  
A:Title: Salivary cystatin SA-III, a potential precursor of the acquired enamel pellicle  
A:Reference number: S16500; MUID:91378515; PMID:1898055  
A:Accession: S16500  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 21-134 'D', 136-141 <IHU>  
R:Isemura, S.; Saitoh, E.; Sanada, K.  
J. Biochem. 96, 489-498, 1984  
A:Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole sal  
A:Reference number: A91985; MUID:85054716; PMID:6501254  
A:Accession: A01272  
A:Molecule type: protein



A;Accession: A32732  
A;Molecule type: protein  
A;Residues: 27-76 <BR>  
R;Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.  
Scand. J. Clin. Lab. Invest. 50, 85-93, 1990  
A;Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst  
A;Reference number: A60552; PMID:90193615; PMID:2315647  
A;Accession: A60552  
A;Molecule type: protein  
A;Residues: 27-49, 'XX', 52-64 <OLA>  
A;Note: This protein, purified from cerebrospinal fluid of patients with the autosomal d  
e defective gene is not present in CSF but is found instead in amyloid deposits  
R;Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.  
Biol. Chem. Hoppe-Seyler 371, 575-580, 1990  
A;Title: Different forms of human cystatin C.  
A;Reference number: S10607; PMID:91025625; PMID:2222856  
A;Accession: S10607  
A;Molecule type: protein  
A;Residues: 27-53 <POP>  
A;Experimental source: urine, kidney disease  
A;Note: truncated forms with amino ends at positions 35 and 36 of the precursor were als  
R;Grubb, A.; Lofberg, H.; Barrett, A.J.  
FEBS Lett. 170, 370-374, 1984  
A;Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
A;Reference number: S01462  
A;Contents: annotation; disulfide bonds  
R;Berti, P.J.; Storer, A.C.  
Biochem. J. 302, 411-416, 1994  
A;Title: Local pH-dependent conformational changes leading to proteolytic susceptibility  
A;Reference number: S55305; PMID:94379969; PMID:8092991  
A;Accession: S55305  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 27-49, 106-146 <BER>  
C;Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl  
I patients with certain autoimmune diseases.  
C;Comment: This protein is an inhibitor of cysteine proteinases and may serve an importa  
C;Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hermorri  
C;Genetics:  
A;Gene: GDB:CST3  
A;Cross-references: GDB:119817; OMIM:105150  
A;Map position: 20p11.2-20p11.2  
A;Introns: 81/3; 119/3  
C;Superfamily: cystatin; cystatin homology  
C;Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline  
F;1-26/Domain: signal sequence #status predicted <SIG>  
F;12-146/Product: cystatin C #status experimental <MAT>  
F;35-146/Domain: cystatin homology <CVS>  
F;81-85/Region: inhibitory #status predicted  
F;29/Modified site: hydroxyproline (Pro) (partial) #status experimental  
F;99-109,123-143/Disulfide bonds: #status experimental

Query Match 36.9%; Score 69; DB 1; Length 146;  
Best Local Similarity 35.5%; Pred. No. 0.051;  
Matches 11; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

Qy 3 SLOWITQYNKESDDKYHFRIFRVLKVRQV 33  
: : : : ||| | : ||| : ||| : ||| :  
Db 52 ALDFAVGEYKNASNDMYHSRALQVVRRKKOI 82

RESULT 6  
S10587  
Cystatin C - rat  
C;Species: Rattus sp. (rat)  
C;Date: 21-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 16-Jul-1999  
C;Accession: S10587  
R;Ennard, F.; Ennard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Brillard, M.; Gauthi  
Biol. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990  
A;Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylat  
A;Reference number: S10587; PMID:90380276; PMID:2400577  
A;Accession: S10587  
A;Status: preliminary



Result No.	Score	Query Match	Length	DB	ID	Description	
1	77	41.2	139	2	A45361	cystatin-related e	
2	74	39.6	141	2	B29632	cystatin SA precu	
3	73	39.0	139	1	UDCH	cystatin precursor	
4	70	37.4	112	1	UDBO	cystatin - bovine	
5	69	36.9	146	1	UDHU	cystatin C precurs	
6	67	35.8	120	2	S10587	cystatin C - rat	
7	67	35.8	137	2	S07085	cystatin C precurs	
8	67	35.8	140	2	A36163	cystatin C precurs	
9	66	35.3	47	2	S62326	cystatin C, paroti	
10	61	32.6	141	2	JQ1470	cystatin S precurs	
11	60	32.1	111	1	JC2040	cystatin - chum sa	
12	60	32.1	132	2	JC4918	cystatin precursor	
13	59	31.6	139	2	E64005	hypothetical prote	
14	58	31.0	141	1	UDHUP1	cystatin S precurs	
15	58	31.0	141	1	UDHUP2	cystatin SN precu	
16	56	29.9	96	2	T07822	cysteine proteinase	
17	56	29.9	162	2	A43428	onchocystatin - ne	
18	54	28.9	235	1	VHVULV	nucleoprotein N -	
19	54	28.9	450	2	A71678	probable response	
20	54	28.9	518	2	T00398	hypothetical prote	
21	53	28.3	111	2	A28793	cystatin - puff ad	
22	53	28.3	132	2	A32239	mercuric resistanc	
23	53	28.3	132	2	T45509	regulatory protein	
24	53	28.3	132	2	T44501	merR1 protein [imp	
25	53	28.3	205	2	T33117	hypothetical prote	
26	53	28.3	319	2	A86650	rhannosyltransfera	
27	52.5	28.1	255	2	A39195	DAL82 protein - ye	
28	52.5	28.1	736	2	S57961	dimethylamine dehy	
29	52	27.8	272	2	T27401	hypothetical prote	

**THIS PAGE BLANK (USPTO)**

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSLOWITDOYNKESDDK 18  
 Db 19 KDSLOWITDOYNKESDDK 36

RESULT 14

US-10-264-049-2608  
 ; Sequence 2608, Application US/10264049  
 ; Publication No. US20040005579A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Birse et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: P133P1  
 ; CURRENT APPLICATION NUMBER: US/10/264,049  
 ; CURRENT FILING DATE: 2002-10-04  
 ; PRIOR APPLICATION NUMBER: PCT/US01/18569  
 ; PRIOR FILING DATE: 2001-06-07  
 ; PRIOR APPLICATION NUMBER: US 60/209,467  
 ; PRIOR FILING DATE: 2000-06-07  
 ; NUMBER OF SEQ ID NOS: 4360  
 ; SOFTWARE: PatentIn Ver. 3.1  
 ; SEQ ID NO 2608  
 ; LENGTH: 181  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-264-049-2608

Query Match 42.2%; Score 79; DB 15; Length 181;  
 Best Local Similarity 42.4%; Pred. No. 0.0051;  
 Matches 14; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1 KDSLOWITDOYNKESDDKXHFIRFVLKVRQV 33  
 Db 85 QOALNEALSEYNKASNDAYHSRAMRVVRVKQV 117

RESULT 15

US-09-775-932-8  
 ; Sequence 8, Application US/09775932  
 ; Patent No. US20020137671A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: University of British Columbia  
 ; TITLE OF INVENTION: Production and use of Modified Cystatins  
 ; FILE REFERENCE: 58069  
 ; CURRENT APPLICATION NUMBER: US/09/775,932  
 ; CURRENT FILING DATE: 2001-02-02  
 ; PRIOR APPLICATION NUMBER: CA99/00717  
 ; PRIOR FILING DATE: 1998-08-05  
 ; PRIOR APPLICATION NUMBER: 60/095,503  
 ; PRIOR FILING DATE: 1998-08-05  
 ; NUMBER OF SEQ ID NOS: 32  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 8  
 ; LENGTH: 121  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-775-932-8

Query Match 39.6%; Score 74; DB 9; Length 121;  
 Best Local Similarity 35.5%; Pred. No. 0.016;  
 Matches 11; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY 3 SLOWITDOYNKESDDKXHFIRFVLKVRQV 33  
 Db 27 ALHFVISEYNKATEDEYRRLRLVLRREQI 57

Search completed: March 18, 2004, 14:30:53  
 Job time : 21.1806 secs

Query Match 52.9%; . Score 99; DB 9; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1  
 OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.5  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2  
 OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUE 5.00e-23  
 OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUE 1.00e-01  
 US-09-864-761-34822

Query Match 92.0%; Score 172; DB 9; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-16;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQ 32  
 |||||  
 Db 19 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQ 50

RESULT 6  
 US-09-864-761-48936  
 ; Sequence 48936, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharon G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 ; FILE REFERENCE: Aeonica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 48936  
 ; LENGTH: 50  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

FEATURE:  
 OTHER INFORMATION: MAP TO AL096677.18  
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96  
 OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUE 5.00e-23  
 OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUE 1.00e-01  
 US-09-864-761-48936

Query Match 92.0%; Score 172; DB 9; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-16;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQ 32  
 |||||  
 Db 19 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQ 50

RESULT 7  
 US-09-941-314-9  
 ; Sequence 9, Application US/09941314  
 ; Patent No. US20020142396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZymoGenetics, Inc.  
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
 ; FILE REFERENCE: 00-81PC  
 ; CURRENT APPLICATION NUMBER: US/09/941,314  
 ; CURRENT FILING DATE: 2001-08-29  
 ; PRIOR APPLICATION NUMBER: 60/230,230  
 ; PRIOR FILING DATE: 2001-09-01  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 27  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-941-314-9

Query Match 76.5%; Score 143; DB 9; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-13;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DQYNKESDDKYHFRIFRVLKVRQVTD 35  
 |||||  
 Db 1 DQYNKESDDKYHFRIFRVLKVRQVTD 27

RESULT 8  
 US-09-941-314-13  
 ; Sequence 13, Application US/09941314  
 ; Patent No. US20020142396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZymoGenetics, Inc.  
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
 ; FILE REFERENCE: 00-81PC  
 ; CURRENT APPLICATION NUMBER: US/09/941,314  
 ; CURRENT FILING DATE: 2001-08-29  
 ; PRIOR APPLICATION NUMBER: 60/230,230  
 ; PRIOR FILING DATE: 2001-09-01  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13  
 ; LENGTH: 49  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-941-314-13

Query Match 76.5%; Score 143; DB 9; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-12;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DQYNKESDDKYHFRIFRVLKVRQVTD 35  
 |||||

FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 115  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-3

Query Match 100.0%; Score 187; DB 9; Length 115;  
Best Local Similarity 100.0%; Pred. No. 3.9e-18;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35  
Db 22 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 56

## RESULT 3

US-09-941-314-4  
Sequence 4, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 117  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-4

Query Match 100.0%; Score 187; DB 9; Length 117;  
Best Local Similarity 100.0%; Pred. No. 4e-18;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35  
Db 24 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 58

## RESULT 4

US-09-941-314-2  
Sequence 2, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 137  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-2

Query Match 100.0%; Score 187; DB 9; Length 137;  
Best Local Similarity 100.0%; Pred. No. 4.7e-18;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35  
Db 44 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 78

## RESULT 5

US-09-864-761-34822  
Sequence 34822, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 34822  
LENGTH: 50  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL109954.10  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 21.1806 Seconds  
(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-8  
Perfect score: 187  
Sequence: 1 KDSLOWITDOYNKESDDKHFRIFRVLKQVQRTD 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	187	100.0	35	9 US-09-941-314-8	Sequence 8, Appli
2	187	100.0	115	9 US-09-941-314-3	Sequence 3, Appli
3	187	100.0	117	9 US-09-941-314-4	Sequence 4, Appli
4	187	100.0	137	9 US-09-941-314-2	Sequence 2, Appli
5	172	92.0	50	9 US-09-864-761-34822	Sequence 34822, A
6	172	92.0	50	9 US-09-864-761-48936	Sequence 48936, A
7	143	76.5	27	9 US-09-941-314-9	Sequence 9, Appli
8	143	76.5	49	9 US-09-941-314-13	Sequence 13, Appli
9	137	73.3	80	9 US-09-941-314-15	Sequence 15, Appli
10	125	66.8	46	9 US-09-941-314-10	Sequence 10, Appli
11	119	63.6	52	9 US-09-941-314-14	Sequence 14, Appli
12	99	52.9	18	9 US-09-941-314-7	Sequence 7, Appli
13	99	52.9	36	9 US-09-941-314-6	Sequence 6, Appli
14	79	42.2	181	15 US-10-264-049-2608	Sequence 2608, Ap
15	74	39.6	121	9 US-09-775-932-8	Sequence 8, Appli

Sequence 24, Appli  
Sequence 6, Appli  
Sequence 16, Appli  
Sequence 15, Appli  
Sequence 4, Appli  
Sequence 16, Appli  
Sequence 24, Appli  
Sequence 2, Appli  
Sequence 17, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 47, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 18, Appli  
Sequence 18, Appli  
Sequence 18, Appli  
Sequence 46, Appli  
Sequence 48, Appli  
Sequence 14, Appli  
Sequence 23, Appli  
Sequence 20, Appli  
Sequence 22, Appli  
Sequence 4, Appli  
Sequence 6, Appli  
Sequence 21, Appli  
Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-09-941-314-8  
; Sequence 8, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-8

Query Match 100.0%; Score 187; DB 9; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.1e-18;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KDSLOWITDOYNKESDDKHFRIFRVLKQVQRTD 35  
Db 1 KDSLOWITDOYNKESDDKHFRIFRVLKQVQRTD 35  
RESULT 2  
US-09-941-314-3  
; Sequence 3, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein

**THIS PAGE BLANK (USPTO)**

```
/ FILING DATE: Filed Herewith
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0193 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 139 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 118195
/ US-08-791-522-4

Query Match 39.0%; Score 73; DB 2; Length 139;
Best Local Similarity 41.9%; Pred. No. 0.0027;
Matches 13; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFIRVLKVRQV 33
Db 47 ALQFAMAENRASNDKYSSRVVRVISAKROL 77

RESULT 14
US-09-314-777-4
/ Sequence 4, Application US/09314777
/ Patent No. 6110686
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Goli, Surya K.
/ TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
/ TITLE OF INVENTION: PROTEIN
/ NUMBER OF SEQUENCES: 4
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/314,777
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/791,522
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0193 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 139 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
```

```
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 118195
/ US-09-314-777-4

Query Match 39.0%; Score 73; DB 3; Length 139;
Best Local Similarity 41.9%; Pred. No. 0.0027;
Matches 13; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFIRVLKVRQV 33
Db 47 ALQFAMAENRASNDKYSSRVVRVISAKROL 77

RESULT 15
US-08-849-303-15
/ Sequence 15, Application US/08849303
/ Patent No. 6680424
/ GENERAL INFORMATION:
/ APPLICANT: Atkinson, Howard J.
/ APPLICANT: McPherson, Michael J.
/ APPLICANT: Urwin, Peter E.
/ TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
/ NUMBER OF SEQUENCES: 79
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Klauber & Jackson
/ STREET: 411 Hackensack Avenue, 4th Floor
/ CITY: Hackensack
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07601
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/849,303
/ FILING DATE: 21-MAY-1997
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Jackson Esq., David A.
/ REGISTRATION NUMBER: 26,742
/ REFERENCE/DOCKET NUMBER: 1321-1-003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 201-487-5800
/ TELEFAX: 201-343-1684
/ TELEX: 133521
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 139 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ US-08-849-303-15

Query Match 39.0%; Score 73; DB 4; Length 139;
Best Local Similarity 41.9%; Pred. No. 0.0027;
Matches 13; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFIRVLKVRQV 33
Db 47 ALQFAMAENRASNDKYSSRVVRVISAKROL 77

Search completed: March 18, 2004, 14:25:36
Job time : 9.95833 secs
```



Query Match 39.6%; Score 74; DB 4; Length 141;  
Best Local Similarity 35.5%; Pred. No. 0.002;  
Matches 11; Conservative 12; Mismatches 8; Indels

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-2

Query Match      44.4%; Score 83; DB 3; Length 141;
Best Local Similarity 45.5%; Pred. No. 9.4e-05;
Matches 15; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRFRVLKVRQVTD 35
Db 47 ALWYAMKEYNKASNDLYNFRVVDILKSQEQITD 79

RESULT 3
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-4

Query Match      43.9%; Score 82; DB 3; Length 142;
Best Local Similarity 45.7%; Pred. No. 0.00013;
Matches 16; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 KDSLQWITDQYNKESDDKYHFRFRVLKVRQVTD 35
Db 46 KQCVWFAMQEYNKESDDKYVFLVVKTLQAQLQVTN 80

RESULT 4
US-09-617-302-4
; Sequence 4, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-4

Query Match      41.2%; Score 77; DB 3; Length 142;
Best Local Similarity 40.0%; Pred. No. 0.00072;
Matches 14; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KDSLQWITDQYNKESDDKYHFRFRVLKVRQVTD 35
Db 46 KQCVWFAMQEYNKESDDKYVFLVVKTLQAQLQVTN 80

RESULT 5
US-09-431-480-3
; Sequence 3, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-431-480-3

Query Match      41.2%; Score 77; DB 3; Length 142;
Best Local Similarity 40.0%; Pred. No. 0.00072;
Matches 14; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KDSLQWITDQYNKESDDKYHFRFRVLKVRQVTD 35
Db 46 KQCVWFAMQEYNKESDDKYVFLVVKTLQAQLQVTN 80

RESULT 6
US-09-617-302-3
; Sequence 3, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-617-302-3

Query Match      41.2%; Score 77; DB 3; Length 142;
Best Local Similarity 40.0%; Pred. No. 0.00072;
Matches 14; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KDSLQWITDQYNKESDDKYHFRFRVLKVRQVTD 35
Db 46 KQCVWFAMQEYNKESDDKYVFLVVKTLQAQLQVTN 80
```



Matches 10; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLOWITDQYNKESDD 35  
 DB 595 LESDEREKIETAVKDALEWLDNQSAEKED 624

RESULT 15

Q9FSY7 PRELIMINARY; PRT; 668 AA.  
 AC Q9FSY7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative luminal binding protein.  
 GN BIP.  
 OS Corylus avellana (European hazel).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fagales; Betulaceae; Corylus.  
 OX NCBI\_TaxID=13451;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RA Gruen S., Suphioglu C., Volkmann D.;  
 RT "Molecular cloning and characterisation of a novel hazel pollen  
 RT allergen identified as a luminal binding protein (Bip).";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF  
 CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 DR EMBL; AJ295617; CAC14168.1; -.  
 DR HSSP; P19120; 3HSC.  
 DR GO; GO:0005783; C:endoplasmic reticulum; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR InterPro; IPR000886; ER\_target\_S.  
 DR InterPro; IPR001023; Hsp70.  
 DR Pfam; PF00012; HSP70; 1.  
 DR PRINTS; PR00301; HEATSHOCK70.  
 DR ProDom; PD000089; Hsp70; 1.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 DR PROSITE; PS00297; HSP70\_1; 1.  
 DR PROSITE; PS00329; HSP70\_2; 1.  
 DR PROSITE; PS01036; HSP70\_3; 1.  
 KW ATP-binding; Endoplasmic reticulum.  
 SQ SEQUENCE 668 AA; 73564 MW; 8243BA3FCF9C10D9 CRC64;

Query Match 29.1%; Score 55; DB 10; Length 668;  
 Best Local Similarity 36.4%; Pred. No. 90;  
 Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 14 VENYAKDSLOWITDQYNKESDD 35  
 DB 603 IESAVKDALEWLDNQSAEKED 624

Search completed: March 18, 2004, 14:21:26  
 Job time : 22.5714 secs

Query Match	29.1%
Best Local Similarity	33.3%

Query Match 29.1%; Score 55; DB 10; Length 665;  
Best Local Similarity 33.3%; Pred. No. 89;

ENBL; AE002178; AAF38049.1; -;  
 ENBL; AP002547; BAA98780.1; -;  
 ENBL; AE017159; AAF98525.1; -;  
 PIR; B86562; B86562.  
 PIR; D72061; D72061.  
 TIGR; CP0175; -;  
 GO; GO:0008237; F-metalloproteinase activity; IEA.  
 GO; GO:0008270; Zinc ion binding; IEA.  
 GO; GO:0006508; Proteolysis and peptidolysis; IEA.  
 InterPro; IPR006025; Pept M Zn BS.  
 PROSITE; PS00142; ZINC PROTEASE; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 184 AA; 22024 MW; EF3682DD1A97A814 CRC64;  
 Query Match 30.4%; Score 57.5; DB 16; Length 184;  
 Best Local Similarity 38.5%; Pred. No. 10;  
 Matches 15; Conservative 6; Mismatches 11; Indels 7; Gaps 1;  
 QY 1 RKTFLSVHEVMAVENYAKD-----SLQWITDQYNKE 32  
 DB 26 RKKTFLSHRVLARPPFPVPCGKGYDLQEIYELNAQ 64  
 RESULT 9  
 O62173 PRELIMINARY; PRT; 209 AA.  
 ID O62173  
 AC O62173  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE F15D3.6 protein.  
 GN F15D3.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA White S.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA none;  
 RA MEDLINE=99069613; PubMed=9851916;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 investigating biology.";  
 RL Science 282:2012-2018 (1998).  
 DR ENBL; Z81063; CAB02955.1; -;  
 DR PIR; T20975; T20975.  
 DR WormPep; F15D3.6; CE15853.  
 DR InterPro; IPR006797; MSFL.  
 DR Pfam; PF04707; MSFL; 1.  
 DR PROSITE; PS50904; PRELI MSFL; 1.  
 SQ SEQUENCE 209 AA; 23667 MW; 346AB71D4BBD39C1 CRC64;  
 Query Match 29.6%; Score 56; DB 5; Length 209;  
 Best Local Similarity 35.5%; Pred. No. 18;  
 Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;  
 QY 2 KKTFLSVHEVMAVENYAKDLSQWITDQYNKE 32  
 DB 140 EKTFLSIYSQNA--NKGQGVVVIDHLKKE 168  
 RESULT 10  
 Q81JTB PRELIMINARY; PRT; 175 AA.  
 ID Q81JTB  
 AC Q81JTB  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 GN RNA polymerase sigma-70 factor, ECF subfamily.  
 EN BA5610.  
 OS Bacillus anthracis (strain Ames).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=198094;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22608414; PubMed=12721629;  
 RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,  
 RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,  
 RA Holtzapple E.K., Okstad O.A., Helgason E., Ristone J., Wu M.,  
 RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,  
 RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,  
 RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,  
 RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,  
 RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,  
 RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,  
 RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolisto A.-B.,  
 RA Fraser C.M.;  
 RT "The genome sequence of Bacillus anthracis Ames and comparison to  
 RT closely related bacteria";  
 RL Nature 423:81-86 (2003).  
 DR ENBL; AE017041; AAP29248.1; -;  
 DR TIGR; BA5610; -;  
 DR InterPro; IPR009043; RNA pol sigma.  
 DR InterPro; IPR007627; Sigma70\_r2.  
 DR InterPro; IPR007630; Sigma70\_r4.  
 DR Pfam; PF04542; sigma70\_r2; 1.  
 DR Pfam; PF04545; sigma70\_r4; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 175 AA; 21395 MW; 58583126B657D80 CRC64;  
 Query Match 29.1%; Score 55; DB 16; Length 175;  
 Best Local Similarity 39.5%; Pred. No. 20;  
 Matches 15; Conservative 4; Mismatches 7; Indels 12; Gaps 2;  
 QY 3 KTFLSVHEVMAVENYAKDLSQWIT-----DOYNKE 32  
 DB 40 KVFSSRHQLSVENYKK-----WITTCVTFYDFYSKK 73  
 RESULT 11  
 O31093 PRELIMINARY; PRT; 327 AA.  
 ID O31093  
 AC O31093  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE PSSG.  
 GN PSSG.  
 OS Rhizobium leguminosarum (biovar viciae).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
 OX NCBI\_TaxID=387;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF39;  
 RA Ivashina T.V., Sadykov M.R., Kanapin A.A., Ksenzenko V.N.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF39;  
 RX MEDLINE=99113394; PubMed=9914965;  
 RA Sadykov M.R., Ivashina T.V., Kanapin A.A., Shlyapnikov M.G.,  
 RA Ksenzenko V.N.;  
 RT "Structural and functional organization of the exopolysaccharide  
 RT biosynthesis genes in Rhizobium leguminosarum bv. viciae VF39";  
 RL Mol. Biol. (Mosk) 32:797-804 (1998).  
 DR ENBL; AF028810; AAB88894.1; -;  
 DR InterPro; IPR001173; Glyco\_trans\_2.  
 DR Pfam; PF00535; Glycos transf 2; 1.  
 SQ SEQUENCE 327 AA; 36361 MW; E3AE2231DCF05A92 CRC64;  
 Query Match 29.1%; Score 55; DB 2; Length 327;  
 Best Local Similarity 47.8%; Pred. No. 41;

DR SMART; SMO0421; HTH\_LUXR; 1.  
 DR PROSITE; PS00622; HTH\_LUXR\_FAMILY; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 240 AA; 28117 MW; 1434A0B458CC54B0 CRC64;

Query Match 30.7%; Score 58; DB 16; Length 240;  
 Best Local Similarity 34.5%; Pred.No.12;  
 Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 3 KTFSLVHEVMVAENYAKDSLQWITDQYNK 31  
 |||:::||::|||:  
 DB 74 KNFLADPVLNPFNSQCHLMWDLLENE 102  
 |||:::|||:

RESULT 6

ID	Q81225	PRELIMINARY;	PRT;	851 AA.
AC	Q81225			
DT	01-MAR-2003	(TrEMBLrel. 23, Created)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)		
DE	Hypothetical protein (Fragment).			
GN	PPA0655W.			
OS	Plasmodium falciparum (isolate 3D7).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
NB	NCBI_TaxID=36329;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22255708; PubMed=12368867;			
RA	Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D., Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K., Bucke C.O., Burrows C., Cherevach I., Chillingworth C., Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C., Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J., Feitwell T., Gobie A., Goodhead I., Gwilliam R., Hamlin N., Hance Z., Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P., Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A., Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N., Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L., Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E., Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M., Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K., Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J., Ralsoulton J.E., Craig A., Newbold C., Barrell B.G.; "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."; Nature 419:527-531(2002). DR EMBL; AL031746; CAB49077.1; -. KW Hypothetical protein. FT NON TER 851 851 SQ SEQUENCE 851 AA; 105268 MW; 76ABED8FCBE9001 CRC64;			

Query Match 30.7%; Score 58; DB 5; Length 851;  
 Best Local Similarity 29.0%; Pred.No.47;  
 Matches 9; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 2 KKTFLSVHEVMVAENYAKDSLQWITDQYNKE 32  
 :::::|||||:  
 DB 742 QSGMINVDNFNKYDDQKFQWIDEKXD 772  
 |||:::|||:

RESULT 7

ID	Q81D83	PRELIMINARY;	PRT;	2030 AA.
AC	Q81D83			
DT	01-MAR-2003	(TrEMBLrel. 23, Created)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)		
DE	MAL13P1.310 protein.			
GN	MAL13P1.310.			
OS	Plasmodium falciparum (isolate 3D7).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
NB	NCBI_TaxID=36329;			
RN	[1]			



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 21.5714 Seconds  
(without alignments)

526.560 Million cell updates/sec

Title: US-09-941-314-6

Perfect score: 189

Sequence: 1 RKKTFLSVHEVAVENYAKDSLQWITDQYNKESDDK 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*  
1: sp archaea:\*  
2: sp bacteria:\*  
3: sp fungi:\*  
4: sp human:\*  
5: sp invertebrate:\*  
6: sp mammal:\*  
7: sp mhc:\*  
8: sp organelle:\*  
9: sp phage:\*  
10: sp plant:\*  
11: sp rodent:\*  
12: sp virus:\*  
13: sp vertebrate:\*  
14: sp unclassified:\*  
15: sp rvirus:\*  
16: sp bacteriaph:\*  
17: sp archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189	100.0	103	4 Q8WXU5	Q8WXU5 homo sapien
2	189	100.0	138	4 Q8WXU6	Q8WXU6 homo sapien
3	100	52.9	139	11 Q8K5A3	Q8K5A3 rattus norv
4	65	34.4	844	5 Q86S89	Q86S89 polypastror
5	58	30.7	240	16 Q8XBD0	Q8XBD0 escherichia
6	58	30.7	851	5 Q81225	Q81225 plasmodium
7	58	30.7	2030	5 Q81B83	Q81B83 plasmodium
8	57.5	30.4	184	16 Q927X9	Q927X9 chlamydia p
9	56	29.6	209	5 Q62173	Q62173 caenorhabdi
10	55	29.1	175	16 Q81J78	Q81J78 bacillus an
11	55	29.1	327	2 Q31093	Q31093 rhizobium l
12	55	29.1	343	16 Q8UD47	Q8UD47 agrobacteri
13	55	29.1	413	16 Q7VXK1	Q7VXK1 haemophilus
14	55	29.1	665	10 Q9W4E8	Q9W4E8 cucumis sat
15	55	29.1	668	10 Q9FSY7	Q9FSY7 corylus ave
16	54.5	28.8	273	17 Q8Z245	Q8Z245 pyrobaculum

17	54.5	28.8	2226	5	Q97225	Q97225 plasmodium
18	54	28.6	176	16	Q814Q9	Q814Q9 bacillus ce
19	54	28.6	189	16	Q8A4V4	Q8A4V4 bacteroides
20	54	28.6	524	10	Q9SBN2	Q9SBN2 volvox cart
21	53.5	28.3	500	16	Q8A815	Q8A815 bacteroides
22	53	28.0	194	5	Q61819	Q61819 caenorhabdi
23	53	28.0	205	5	Q8T7Y7	Q8T7Y7 caenorhabdi
24	53	28.0	208	16	Q7UAB1	Q7UAB1 shigella fl
25	53	28.0	212	16	Q83R45	Q83R45 shigella fl
26	53	28.0	244	16	Q8FGM5	Q8FGM5 escherichia
27	53	28.0	300	5	Q86KT1	Q86KT1 dictyostell
28	53	28.0	351	2	Q48418	Q48418 klebsiella
29	53	28.0	351	2	Q8GPR0	Q8GPR0 citrobacter
30	53	28.0	543	9	Q859F0	Q859F0 pseudomonas
31	52	27.5	321	2	Q8G8T6	Q8G8T6 pseudomonas
32	52	27.5	404	10	Q9SY99	Q9SY99 arabidopsis
33	52	27.5	457	10	Q9SU04	Q9SU04 arabidopsis
34	52	27.5	592	3	Q8X228	Q8X228 mycosphaere
35	52	27.5	1227	5	Q9BIA2	Q9BIA2 caenorhabdi
36	52	27.5	1422	5	O00914	O00914 plasmodium
37	51.5	27.2	122	15	Q9EAA5	Q9EAA5 human immun
38	51.5	27.2	318	16	Q8X714	Q8X714 escherichia
39	51.5	27.2	465	16	Q9FJ81	Q9FJ81 campylobact
40	51.5	27.2	567	17	O26602	O26602 methanobact
41	51.5	27.2	656	5	Q24842	Q24842 entamoeba h
42	51	27.0	190	16	Q97KN6	Q97KN6 clostridium
43	51	27.0	216	3	Q05435	Q05435 saccharomyc
44	51	27.0	238	16	Q8CN34	Q8CN34 staphylococ
45	51	27.0	326	1	Q9UXP0	Q9UXP0 methanoblob

#### ALIGNMENTS

RESULT 1  
Q8WXU5 ID Q8WXU5 PRELIMINARY; PRT; 103 AA.  
AC Q8WXU5;  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
DE SC13delta.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;  
RT "SC13: A novel epididymal specific member of the cystatin family.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF335481; AAL71992.1; -;  
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
DR InterPro; IPR00010; Cystatin.  
DR Pfam; PF00031; Cystatin; 1.  
SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;

Query Match 100.0%; Score 189; DB 4; Length 103;  
Best Local Similarity 100.0%; Pred. No. 1.4e-17;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKKTFLSVHEVAVENYAKDSLQWITDQYNKESDDK 36  
|||||  
Db 27 RKKTFLSVHEVAVENYAKDSLQWITDQYNKESDDK 62

RESULT 2  
Q8WXU6 ID Q8WXU6 PRELIMINARY; PRT; 138 AA.  
AC Q8WXU6;  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)

```

OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=92361242; PubMed=1822990;
RX Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RA "The tobacco luminal binding protein is encoded by a multigene
RT family.";
RT Plant Cell 3:1025-1035(1991).
RN [2]
RN ERRATUM.
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RL Plant Cell 3:1251-1251(1991).
CC -!- FUNCTION: Probably plays a role in facilitating the assembly of
CC multimeric protein complexes inside the ER.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60059; CAA42661.1; -.
DR PIR; P02622; S21878.
DR HSP; P08109; 1CKR.
DR InterPro; IPR000886; ER_target_S.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR PROSITE; PS00297; HSP70_1; PARTIAL.
DR PROSITE; PS00329; HSP70_2; PARTIAL.
DR PROSITE; PS01036; HSP70_3; PARTIAL.
DR PROSITE; PS00014; ER_TARGET; 1.
KW ATP-binding; Endoplasmic reticulum; Multigene family.
FT NON_TER 1
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 289 292 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 292 AA; 32260 MW; 6E7A4F5107C6E2D5 CRC64;

Query Match 27.0%; Score 51; DB 1; Length 292;
Best Local Similarity 30.0%; Pred. No. 18;
Matches 9; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLOWITDQYNKESDD 35
DB 219 LESDEKEKIEATKAELEWLDNQSAEKD 248

RESULT 15
BIP8 TOBAC
ID BIP8 TOBAC STANDARD; PRT; 293 AA.
AC Q03686;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luminal binding protein 8 (Bip 8) (78 kDa glucose-regulated protein
DE homolog 8) (GRP 78-8) (Fragment).
GN Bip8.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RN SEQUENCE FROM N.A.

```

```

RX MEDLINE=92361242; PubMed=1822990;
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RT "The tobacco luminal binding protein is encoded by a multigene
RT family.";
RT Plant Cell 3:1025-1035(1991).
RN [2]
RN ERRATUM.
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RL Plant Cell 3:1251-1251(1991).
CC -!- FUNCTION: Probably plays a role in facilitating the assembly of
CC multimeric protein complexes inside the ER.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60062; CAA42664.1; -.
DR PIR; S21881; S21881.
DR HSP; P08109; 1CKR.
DR InterPro; IPR000886; ER_target_S.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR PROSITE; PS00297; HSP70_1; PARTIAL.
DR PROSITE; PS00329; HSP70_2; PARTIAL.
DR PROSITE; PS01036; HSP70_3; PARTIAL.
DR PROSITE; PS00014; ER_TARGET; 1.
KW ATP-binding; Endoplasmic reticulum; Multigene family.
FT NON_TER 1
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 290 293 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 293 AA; 32342 MW; 44939595E45650E6 CRC64;

Query Match 27.0%; Score 51; DB 1; Length 293;
Best Local Similarity 30.0%; Pred. No. 18;
Matches 9; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLOWITDQYNKESDD 35
DB 222 LESDEKEKIEATKAELEWLDNQSAEKD 251

Search completed: March 18, 2004, 14:16:10
Job time : 5.28571 secs

```

```

QY 1 RKTFSLVHEVMAVENYAKDSLOWITDQYNKESDD 35
   : : : : : : : : : : : : : : : : : :
Db 52 RTLYLSVEE---QDNLOEALIRWLCYHNLNEDD 83

RESULT 12
BIP1_TOBAC
ID_BIP3_TOBAC STANDARD; PRT; 168 AA.
AC Q03683;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luminal binding protein 3 (BIP 3) (78 kDa glucose-regulated protein
DE homolog 3) (GRP 78-3) (Fragment).
GN BIP3.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92361242; PubMed=1822990;
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RT "The tobacco luminal binding protein is encoded by a multigene
RT family.";
RL Plant Cell 3:1025-1035(1991).
RN [2]
RP ERRATUM.
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RL Plant Cell 3:1251-1251(1991).
CC -!- FUNCTION: Probably plays a role in facilitating the assembly of
CC multimeric protein complexes inside the ER.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X60061; CAA42663.1; -
CC HSP; P04475; 1DKY.
CC InterPro; IPR000886; ER target_S.
CC Pfam; PF00012; Hsp70.
CC ProDom; PD000089; Hsp70; 1.
CC PROSITE; PS00297; HSP70_1; PARTIAL.
CC PROSITE; PS00329; HSP70_2; PARTIAL.
CC PROSITE; PS00336; HSP70_3; PARTIAL.
CC PROSITE; PS00014; ER TARGET; 1.
CC ATP-binding; Endoplasmic reticulum; Multigene family.
CC NON_TER 1
CC CARBOHYD 120 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SITE 165 168 PREVENT SECRETION FROM ER (POTENTIAL).
CC SEQUENCE 168 AA; 18838 MW; D0089CF2219C624E CRC64;

Query Match 27.0%; Score 51; DB 1; Length 168;
Best Local Similarity 30.0%; Pred. No. 9.6;
Matches 9; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLOWITDQYNKESDD 35
   : : : : : : : : : : : : : : : : : :
Db 98 LESDEKEKETATKEALEWLDNQSAEKED 127

RESULT 13
BIP1_TOBAC
ID_BIP1_TOBAC STANDARD; PRT; 290 AA.
AC Q03682;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luminal binding protein 2 (BIP 2) (78 kDa glucose-regulated protein
DE homolog 2) (GRP 78-2) (Fragment).
GN BIP2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92361242; PubMed=1822990;
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RT "The tobacco luminal binding protein is encoded by a multigene
RT family.";
RL Plant Cell 3:1025-1035(1991).
RN [2]
RP ERRATUM.
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RL Plant Cell 3:1251-1251(1991).
CC -!- FUNCTION: Probably plays a role in facilitating the assembly of
CC multimeric protein complexes inside the ER.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X60060; CAA42662.1; -
CC PIR; S21877; S21877.
CC HSP; P08109; 1CKR.
CC InterPro; IPR000886; ER target_S.
CC InterPro; IPR001023; Hsp70.
CC Pfam; PF00012; Hsp70; 1.
CC ProDom; PD000089; Hsp70; 1.
CC PROSITE; PS00297; HSP70_1; PARTIAL.
CC PROSITE; PS00329; HSP70_2; PARTIAL.
CC PROSITE; PS00336; HSP70_3; PARTIAL.
CC PROSITE; PS00014; ER TARGET; 1.
CC ATP-binding; Endoplasmic reticulum; Multigene family.
CC NON_TER 1
CC CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SITE 287 290 PREVENT SECRETION FROM ER (POTENTIAL).
CC SEQUENCE 290 AA; 32031 MW; E80F65136B64BP89 CRC64;

Query Match 27.0%; Score 51; DB 1; Length 290;
Best Local Similarity 30.0%; Pred. No. 18;
Matches 9; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLOWITDQYNKESDD 35
   : : : : : : : : : : : : : : : : : :
Db 219 LESDEKEKETATKEALEWLDNQSAEKED 248

RESULT 14
BIP2_TOBAC
ID_BIP2_TOBAC STANDARD; PRT; 292 AA.
AC Q03682;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luminal binding protein 2 (BIP 2) (78 kDa glucose-regulated protein
DE homolog 2) (GRP 78-2) (Fragment).
GN BIP2.

```

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN NCBI\_TaxID=5860;  
 [1]  
 RX SEQUENCE FROM N.A.  
 RA MEDLINE=95059225; PubMed=7969277;  
 Cheng O., Saul A.;  
 RT "the dihydrofolate reductase domain of rodent malarial: point  
 mutations and pyrimethamine resistance.";  
 RL Mol. Biochem. Parasitol. 5; 6:361-363(1994).  
 CC -1- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADP(+) = 7,8-  
 dihydrofolate + NADPH;  
 CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP =  
 dihydrofolate + dUMP;  
 CC -1- PATHWAY: Essential step for de novo glycine and purine synthesis,  
 DNA precursor synthesis, and for the conversion of dUMP to dTMP.  
 CC -1- SIMILARITY: In the N-terminal section; belongs to the  
 dihydrofolate reductase family.  
 CC -1- SIMILARITY: In the C-terminal section; belongs to the thymidylate  
 synthase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L28121; AAA29582.1; -;  
 DR HSSP; P00379; 1RA9.  
 DR InterPro; IPR001796; DHFR.  
 DR ProSite; PS000398; Thymidylat\_synth.  
 DR Pfam; PF00186; Dihfolate\_red; 1.  
 DR PRINTS; PR00070; DHFR.  
 DR PROSITE; PS00075; DHFR; 1.  
 DR PROSITE; PS00091; THYMIDYLATE SYNTHASE; FALSE\_NEG.  
 KW Multifunctional enzyme; Oxidoreductase; Transferase; NADP;  
 KW Methyltransferase; Nucleotide biosynthesis; One-carbon metabolism.  
 FT NON TER 1  
 FT DOMAIN <1> 182 DIHYDROFOLATE REDUCTASE.  
 FT NON TER 182  
 FT NON TER 182  
 SQ SEQUENCE 182 AA; 20924 MW; 33CD27441763198 CRC64;  
 Query Match 27.5%; Score 52; DB 1; Length 182;  
 Best Local Similarity 38.2%; Pred. No. 7.7;  
 Matches 13; Conservative 7; Mismatches 12; Indels 2; Gaps 1;  
 QY 3 KTFSLVHEVMAVENYAKDSLQWTDQYNKESDDK 36  
 DB 42 KHFVSVTSYNNENYIR--LKWKRDKYIKENNVK 73  
 RESULT 11  
 HEMI\_SALTY STANDARD; PRT; 418 AA.  
 AC P13581;  
 ID 01-JAN-1990 (Rel. 13, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Glutamyl-tRNA reductase (EC 1.2.1.1-) (GLUTR).  
 GN HEMA OR STM1777 OR STY1902 OR T1099.  
 OS Salmonella typhimurium, and  
 OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OC NCBI\_TaxID=602, 601;  
 RN [1]  
 RX SEQUENCE FROM N.A.  
 RA MEDLINE=89291746; PubMed=2544564;  
 Elliott T.;  
 RT "Cloning, Genetic characterization, and nucleotide sequence of the  
 hemA-prfA operon of Salmonella typhimurium.";

RL J. Bacteriol. 171:3948-3960(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 Waterston R., Wilson R.K.;  
 RA "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 LT2.";  
 RT Nature 413:852-856(2001).  
 RL [3]  
 RN SEQUENCE FROM N.A.  
 RP SPECIES=S.typhi; STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahinia M.,  
 Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 Whitehead S., Barrrell B.G.;  
 RA "Complete genome sequence of a multiple drug resistant Salmonella  
 enterica serovar Typhi CT18.";  
 RT Nature 413:848-852(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;  
 MEDLINE=22531367; PubMed=12644504;  
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
 Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;  
 RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
 and CT18.";  
 RT J. Bacteriol. 185:2330-2337(2003).  
 CC -1- CATALYTIC ACTIVITY: Glutamyl-tRNA (Glu) + NADPH = glutamate-1-  
 semialdehyde + NADP(+) + tRNA (Glu).  
 CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.  
 CC -1- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J04243; AAA88610.1; -;  
 DR EMBL; A8008779; AAL20692.1; -;  
 DR EMBL; AL627271; CAD02132.1; -;  
 DR EMBL; A8016837; AAC06762.1; -;  
 DR PIR; A32661; BVEBHA.  
 DR HSSP; Q42843; 1B29.  
 DR HAMAP; SG10149; hemA.  
 DR StyGene; MF 00087; -; 1.  
 DR InterPro; IPR000343; Glutr.  
 DR Pfam; PF00745; Glutr\_dimer; 1.  
 DR Pfam; PF05201; Glutr\_N; 1.  
 DR Pfam; PF05200; Glutr\_NAD\_bind; 1.  
 DR TIGRFAMs; TIGR01035; hemA; 1.  
 DR PROSITE; PS00747; GLUTR; 1.  
 DR Porphyrin biosynthesis; Oxidoreductase; NADP; Complete proteome.  
 KW ACT\_SITE 50 50 NUCLEOPHILE (BY SIMILARITY).  
 FT ACT\_SITE 99 99 BASE (BY SIMILARITY).  
 FT CONFLICT 139 139 E -> R (IN REF. 1).  
 SQ SEQUENCE 418 AA; 46105 MW; 92961964C36AD3F4 CRC64;  
 Query Match 27.2%; Score 51.5; DB 1; Length 418;  
 Best Local Similarity 31.4%; Pred. No. 23;  
 Matches 11; Conservative 8; Mismatches 13; Indels 3; Gaps 1;

```

RP SEQUENCE FROM N.A.
RX MEDLINE=86176730; PubMed=3515318;
RA Sharma S., Stark T.F., Beattie W.G., Moses R.E.;
RT "Multiple control elements for the uvrC gene unit of Escherichia
RL coli.";
RN Nucleic Acids Res. 14:2301-2318(1986).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251350; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubram S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RL corresponding to the 40.1-50.0 min region on the linkage map.";
RN DNA Res. 3:379-392(1996).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=92007787; PubMed=1915297;
RA Wang X., de Boer P.A.J., Rothfield L.I.;
RT "A factor that positively regulates cell division by activating
RL transcription of the major cluster of essential cell division genes
of Escherichia coli.";
RN EMBO J. 10:3363-3372(1991).
CC -1- FUNCTION: ACTIVATES CELL DIVISION BY SPECIFICALLY INCREASING
CC TRANSCRIPTION FROM ONE OF THE TWO PROMOTERS THAT LIE IMMEDIATELY
CC UPSTREAM OF THE FTSQAZ GENE CLUSTER.
CC -1- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X03691; CAA27327.1; -.
DR EMBL; AE000284; AAC74983.1; -.
DR EMBL; D90832; BAA15736.1; -.
DR PIR; A64955; Q08CUI.
DR EcoGene; EG10935; sdia.
DR InterPro; IPR005143; Autoind_bind.
DR InterPro; IPR000792; HTH_LuxR.
DR Pfam; PF03472; Autoind_bind; 1.
DR Pfam; PF00196; GerE; 1.
DR ProDom; PD000307; HTH_LuxR; 1.
DR SMART; SM00421; HTH_LuxR; 1.
DR PROSITE; PS00622; HTH_LuxR_FAMILY; 1.
KW Transcription regulation; Activator; DNA-binding; Trans-acting factor;
KW Cell division; Complete proteome.
FT DNA BIND 197 216 H-T-H MOTIF (BY SIMILARITY).
FT CONFLICT 120 128 TQVLMPLNR -> HSVFNAAQTG (IN REF. 1).
FT CONFLICT 140 140 A -> R (IN REF. 1).
SQ SEQUENCE 240 AA; 28117 MW; C04CC50C06135C4 CRC64;

Query Match 28.0%; Score 53; DB 1; Length 240;
Best Local Similarity 31.0%; Pred. No. 7.7;
Matches 9; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 3 KITFLSHEVMAVENYAKDSLOWITDQYNK 31
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
Db 74 KNFLAIDPVLNPFNSQGHLMWDDLFSE 102

RESULT 9
PREI STAAU STANDARD; PRT; 403 AA.
AC P03857;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Plasmid recombination enzyme (Mobilization protein).
GN PRE OR MOB.
OS Staphylococcus aureus.
OG Plasmid pE194.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82167187; PubMed=6279574;
RA Horinouchi S., Weisblum B.;
RT "Nucleotide sequence and functional map of pE194, a plasmid that
RT specifies inducible resistance to macrolide, lincosamide, and
RT streptogramin type B antibiotics.";
RL J. Bacteriol. 150:804-814(1982).
CC -1- FUNCTION: THE INTERACTION OF THE RSA SITE AND THE PRE PROTEIN MAY
CC NOT ONLY SERVES A FUNCTION IN PLASMID MAINTENANCE, BUT MAY ALSO
CC CONTRIBUTE TO THE DISTRIBUTION OF SMALL ANTIBIOTIC RESISTANCE
CC PLASMIDS AMONG GRAM-POSITIVE BACTERIA.
CC -1- MISCELLANEOUS: PRE PROTEINS CONTAIN CONSERVED POSITIVELY CHARGED
CC AMINO ACIDS PROBABLY INVOLVED IN THE BINDING OF THE PRE PROTEIN TO
CC THE RSA SITE.
CC -1- SIMILARITY: TO OTHER PRE PROTEINS (FROM PLASMIDS PUB110, PMV158,
CC PE194, PT181, PTB913), IN THEIR N-TERMINAL ONLY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; V01278; CAA24593.1; -.
DR InterPro; IPR001668; Mob_Pre.
DR Pfam; PF01076; Mob_Pre; 1.
DR KMW; DNA-binding.
FT BINDING 45 45 DNA (POTENTIAL).
FT BINDING 115 115 DNA (POTENTIAL).
SQ SEQUENCE 403 AA; 47839 MW; 875867A394000FC1 CRC64;

Query Match 28.0%; Score 53; DB 1; Length 403;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 8; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 14 VENVAKDSLOWITDQYNKES 33
| | | | | | | | | | | | | | |
| | | | | | | | | | | | | | |
Db 100 IERFKDSLEFLENEYCKEN 119

RESULT 10
DRTS PLAVN STANDARD; PRT; 182 AA.
AC P46103;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-TS)
DE [includes: Dihydrofolate reductase (EC 1.5.1.3); Thymidylate synthase
DE (EC 2.1.1.45)] (Fragment).
OS Plasmodium vinckei.

```

```
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 663 LUMINAL BINDING PROTEIN 2.
FT CARBOHYD 614 614 N-LINKED (GLCNAC... ) (POTENTIAL).
FT SITE 660 663 PREVENT SECRETION FROM ER.
SQ SEQUENCE 663 AA; 73084 MW; 676987D19A0D918A CRC64;

Query Match 28.6%; Score 54; DB 1; Length 663;
Best Local Similarity 33.3%; Pred. No. 18;
Matches 10; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLQWITDQYNKESDD 35
DB 592 LEAEKEKVEEALKEALEWLDNDSQAEKED 621

RESULT 6
BIP3_MAIZE STANDARD; PRT; 663 AA.
AC O24581;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luminal binding protein 3 precursor (BiP3).
GN BIPE3.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=cv. Flourey-2; TISSUE=Kernel;
RX MEDLINE=98094260; PubMed=9434171;
RA Wobbel R.L., Obrian G.R., Boston R.S.;
RT "Comparative analysis of BiP gene expression in maize endosperm.";
RL Gene 204:105-113(1997).
CC -1- FUNCTION: Probably plays a role in facilitating the assembly of
CC multimeric protein complexes inside the ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U58209; AAC49900.1; -.
DR PIR; T04080; T04080.
DR HSP; P19120; 3HSC.
DR InterPro; IPR000886; ER_target_S.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; P00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 663 LUMINAL BINDING PROTEIN 3.
FT CARBOHYD 614 614 N-LINKED (GLCNAC... ) (POTENTIAL).
FT SITE 660 663 PREVENT SECRETION FROM ER.
SQ SEQUENCE 663 AA; 73156 MW; A710278D9F692723 CRC64;

Query Match 28.6%; Score 54; DB 1; Length 663;
Best Local Similarity 33.3%; Pred. No. 18;
Matches 10; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLQWITDQYNKESDD 35
DB 592 LEAEKEKVEEALKEALEWLDNDSQAEKED 621
```

```
DB 592 LEAEKEKVEEALKEALEWLDNDSQAEKED 621
MLR_LUMTE STANDARD; PRT; 195 AA.
AC P80164;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myosin regulatory light chain, striated muscle, 25 kDa isoform (LC25).
OS Lumbricus terrestris (Common earthworm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Lumbricus.
OX NCBI_TaxID=6398;
RN [1]
SEQUENCE.
RX MEDLINE=93145966; PubMed=8425543;
RA Serwe M., Meyer H.E., Craig A.G., Carlhoff D., D'Haese J.;
RT "Complete amino acid sequence of the regulatory light chain of
RT obliquely striated muscle myosin from earthworm, Lumbricus
RT terrestris.";
RL Eur. J. Biochem. 211:341-346(1993).
CC -1- FUNCTION: Plays an important role in regulation of muscle cell
CC contractile activity.
CC -1- SUBUNIT: Myosin is an hexamer of 2 heavy chains and 4 light
CC chains.
CC -1- MISCELLANEOUS: This chain binds calcium.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
DR PIR; S28845; S28845.
DR HSP; P13543; 1SCW.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 2.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; EFh; 2.
DR PROSITE; PS00018; EF_HAND; 1.
KW Myosin; Calcium-binding; Muscle protein.
FT DOMAIN 2 14 LYS-RICH (BASIC).
FT CA BIND 68 79 EF-HAND (BY SIMILARITY).
FT VARIANT 65 65 T -> A (IN 20% OF THE MOLECULES).
FT VARIANT 70 70 D -> N (IN 20% OF THE MOLECULES).
FT VARIANT 76 78 GPD -> DES (IN 20% OF THE MOLECULES).
FT VARIANT 81 82 GN -> AA (IN 20% OF THE MOLECULES).
FT VARIANT 84 84 F -> Y (IN 20% OF THE MOLECULES).
FT VARIANT 127 128 GT -> SS (IN 20% OF THE MOLECULES).
FT VARIANT 141 141 L -> V (IN 20% OF THE MOLECULES).
FT VARIANT 168 168 V -> L (IN 20% OF THE MOLECULES).
SQ SEQUENCE 195 AA; 21945 MW; 34BF85EB4D4E7F73 CRC64;

Query Match 28.0%; Score 53; DB 1; Length 195;
Best Local Similarity 52.6%; Pred. No. 6.2;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 14 VENYAKDSLQWITDQYNKE 32
DB 145 LEEYVXDLTLNVGDQYNKD 163

RESULT 8
SDIA_ECOLI STANDARD; PRT; 240 AA.
AC P07026; P76313;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Regulatory protein sdia.
GN SDIA OR B1916.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
```

DR PIR; D64151; D64151.  
DR TIGR; H10404; -.  
DR InterPro; IPR000541; UPF0021.  
DR Pfam; PF01171; ATP\_bind3; 1.  
DR Complete proteome.  
KW SEQUENCE 430 AA; 50052 MW; FFDC831DB67C1C6C CRC64;  
Query Match 31.5%; Score 59.5; DB 1; Length 430;  
Best Local Similarity 52.4%; Pred. No. 2;  
Matches 11; Conservative 6; Mismatches 3; Indels 1; Gaps 1;  
QY 14 VENYA-KDSLOWITDQYNKES 33  
:|||||:|:|||||:|:|:  
Db 169 LENYAQKEKLNWITDSNEDN 189  
RESULT 4  
FR12\_SOYBN STANDARD; PRT; 257 AA.  
AC Q94IC4;  
ID 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ferritin 2, chloroplast precursor (SferH-2).  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
OX NCBI\_TaxID=384;  
RN [1]  
SEQUENCE FROM N.A.  
RP MEDLINE=21276460; PubMed=11278989;  
RX Masuda T., Goto F., Yoshihara T.;  
RA "A novel plant ferritin subunit from soybean that is related to a  
mechanism in iron release.";  
RT J. Biol. Chem. 276:19575-19579(2001).  
RL CC  
-1- FUNCTION: Ferritin is an intracellular molecule that stores iron  
in a soluble, nontoxic, readily available form. The functional  
molecule, which is composed of 24 chains, is roughly spherical and  
contains a central cavity into which the polymeric ferric iron  
core is deposited.  
CC  
-1- SUBCELLULAR LOCATION: Chloroplast.  
CC  
-1- SIMILARITY: Belongs to the ferritin family.  
CC  
-1- SIMILARITY: Contains 1 ferritin-like diiron domain.  
CC  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
-----  
CC EMBL; AB062754; BAB60683.1; -.  
CC InterPro; IPR001519; Ferritin.  
CC InterPro; IPR008331; Ferritin\_Dns.  
CC InterPro; IPR009040; Ferritin\_like.  
CC Pfam; PF00210; ferritin; 1.  
CC ProDom; PD000971; Ferritin; 1.  
CC PROSITE; PS00540; FERRITIN\_1; FALSE\_NRG.  
CC PROSITE; PS00204; FERRITIN\_2; 1.  
CC PROSITE; PS00905; FERRITIN-LIKE; 1.  
CC Iron storage; Iron; Metal-Binding; Chloroplast; Transit peptide.  
FT TRANSIT 1 51 CHLOROPLAST (POTENTIAL).  
FT CHAIN 52 257 FERRITIN 2.  
FT DOMAIN 52 84 EXTENSION PEPTIDE (EP).  
FT DOMAIN 85 238 FERRITIN-LIKE DIIRON.  
FT METAL 102 102 IRON (BY SIMILARITY).  
FT METAL 136 136 IRON (BY SIMILARITY).  
FT METAL 137 137 IRON (BY SIMILARITY).  
FT METAL 139 139 IRON (BY SIMILARITY).  
FT METAL 140 140 IRON (BY SIMILARITY).  
FT METAL 186 186 IRON (BY SIMILARITY).  
FT METAL 186 186 IRON (BY SIMILARITY).  
KW Iron storage; Iron; Metal-Binding; Chloroplast; Transit peptide.  
FT TRANSIT 1 51 CHLOROPLAST (POTENTIAL).  
FT CHAIN 52 257 FERRITIN 2.  
FT DOMAIN 52 84 EXTENSION PEPTIDE (EP).  
FT DOMAIN 85 238 FERRITIN-LIKE DIIRON.  
FT METAL 102 102 IRON (BY SIMILARITY).  
FT METAL 136 136 IRON (BY SIMILARITY).  
FT METAL 137 137 IRON (BY SIMILARITY).  
FT METAL 139 139 IRON (BY SIMILARITY).  
FT METAL 140 140 IRON (BY SIMILARITY).  
FT METAL 186 186 IRON (BY SIMILARITY).  
FT METAL 186 186 IRON (BY SIMILARITY).  
KW Iron storage; Iron; Metal-Binding; Chloroplast; Transit peptide.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 4.28571 Seconds  
(without alignments)  
437.389 Million cell updates/sec

Title: US-09-941-314-6  
Perfect score: 189  
Sequence: 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues  
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189	100.0	137	CS11 HUMAN	Q9h112 homo sapien
2	104	55.0	139	CS11 MOUSE	Q9d269 mus musculus
3	59.5	31.5	430	MESJ HABIN	P44689 haemophilus
4	55	29.1	257	FR12 SOYBN	Q941c4 glycine max
5	54	28.6	663	BIP2 MAIZE	Z24067 zea mays (m
6	54	28.6	663	BIP3 MAIZE	P80164 lumbricus t
7	53	28.0	195	MLR LUMTE	P07026 escherichia
8	53	28.0	240	SDIA ECOLI	P03857 staphylococ
9	53	28.0	403	PRE1 STAAU	P46103 plasmodium
10	52	27.5	182	DRTS PLAVN	FI3581 salmonella
11	51.5	27.2	418	HEMA_SALTY	Q03683 nicotiana t
12	51	27.0	168	BIP3 TOBAC	Q03681 nicotiana t
13	51	27.0	290	BIP1 TOBAC	Q03682 nicotiana t
14	51	27.0	292	BIP2 TOBAC	Q03686 nicotiana t
15	51	27.0	293	BIP8 TOBAC	Q03686 nicotiana t
16	51	27.0	326	MER METTI	Q09x00 methanolobu
17	51	27.0	438	GATD PYRFU	Q8u0X0 pyrococcus
18	51	27.0	438	GATD PYRHO	O59132 pyrococcus
19	51	27.0	666	BIP LYCES	P49118 lycopersico
20	51	27.0	667	BIP4 TOBAC	Q03684 nicotiana t
21	51	27.0	668	BIP5 TOBAC	Q03685 nicotiana t
22	51	27.0	668	BIP SPIOL	Q42434 spinacia ol
23	51	27.0	838	OS94 MOUSE	P48722 mus musculus
24	51	27.0	1818	HMW2 MYCFN	P75471 mycoplasma
25	50	26.5	302	TX34 CAEEL	Q9nah2 caenorhabdi
26	49.5	26.2	418	HEM1 ECOL6	Q8fi03 escherichia
27	49.5	26.2	439	RHGI HUMAN	Q07960 homo sapien
28	49.5	26.2	611	SNF1 CANGA	Q00372 candida gla
29	49.5	26.2	976	SCPI1 HUMAN	Q15431 homo sapien
30	49	25.9	328	MER METWA	Q8p266 methanosarc
31	49	25.9	438	GATD PYRAB	Q9v0t9 pyrococcus
32	49	25.9	474	SYE VIBPA	Q87rl6 vibrio para
33	49	25.9	484	SYE_STAMM	Q99w75 staphylococ

34	49	25.9	657	1	LONI THEAC
35	49	25.9	918	1	IL6B HUMAN
36	48.5	25.7	326	1	FAH2 STAEP
37	48.5	25.7	504	1	GUNW_ERWCA
38	48.5	25.7	792	1	PEPX_LACDL
39	48.5	25.7	793	1	PEPX_LACHE
40	48.5	25.7	944	1	TPSX SCHPO
41	48	25.4	324	1	Y089 BORBU
42	48	25.4	326	1	XC00_PVRAB
43	48	25.4	348	1	VIPB_SALTY
44	48	25.4	426	1	CG2A PATVU
45	48	25.4	444	1	PST1_YEAST

## ALIGNMENTS

RESULT 1  
CS11\_HUMAN STANDARD; PRT; 137 AA.  
AC Q9H112; Q9H113;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE Cystatin 11 precursor.  
GN Cst11 OR CST8L.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,  
Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Corby N.R.,  
Clegg S., Cobley V.E., Collier R.E., Connor R.E., Dunn M.,  
Ellington A.G., Coville G.J., Deadman R., Dhali P.D., Dunn M.,  
Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Garner P.,  
Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
Lehvaeslaiho M.H., Leveraha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A.A.,  
Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
Whithead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
Rogers J.;  
"The DNA sequence and comparative analysis of human chromosome 20.";  
Nature 414:865-871(2001).  
RL -1- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q9H112-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q9H112-2; Sequence=VSP\_001260;  
CC Notes=No experimental confirmation available;  
CC -1- SIMILARITY: Belongs to the cystatin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 XX  
 SQ Sequence 117 AA;  
 Query Match 100.0%; Score 189; DB 5; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-18;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36  
 |||||  
 DB 6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 41  
 |||||

CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)  
 XX  
 SQ Sequence 137 AA;  
 Query Match 100.0%; Score 189; DB 5; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-18;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36  
 |||||  
 DB 26 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 61  
 |||||

Search completed: March 18, 2004, 14:14:58  
 Job time : 33.6429 secs

Search completed: March 18, 2004, 14:14:58  
 Job time : 33.6429 secs

RESULT 15  
 AAU79852  
 ID AAU79852 standard; protein; 137 AA.  
 XX  
 AC AAU79852;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8).  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2002020567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI: 2002-383044/41.  
 DR N-PSDB; ABK49522.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 93-94; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and

CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Rudik syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a peptide/protein encoded by a single exon probe of  
 CC the invention. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 50 AA;  
 Query Match 100.0%; Score 189; DB 5; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1e-18;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36  
 DB 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36  
 RESULT 13  
 AAU79853  
 ID AAU79853 standard; protein; 115 AA.  
 XX  
 AC AAU79853;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #1.  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic fragment.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868. *see in situ*  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 94; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the

CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 XX  
 SQ Sequence 115 AA;  
 Query Match 100.0%; Score 189; DB 5; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-18;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36  
 DB 4 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 39  
 RESULT 14  
 AAU79854  
 ID AAU79854 standard; protein; 117 AA.  
 XX  
 AC AAU79854;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #2.  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic fragment.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 94-95; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36  
|||||

Db 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36  
|||||

## RESULT 11

AA02833  
ID AA02833 standard; protein; 50 AA.  
XX  
AC AA02833;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Peptide #1515 encoded by probe for measuring breast gene expression.  
XX  
KW Probe; human; breast disease; breast cancer; development disorder;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN W0200157270-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US000661.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-476286/51.  
XX  
PT Novel single exon nucleic acid probe used to measuring gene expression in  
PT a human breast.  
XX  
PS Claim 27; SEQ ID NO 11573; 322pp; English.  
XX

CC The present invention relates to novel single exon nucleic acid probes  
CC (see AA00010-AA10067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridises at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosing  
CC diseases of the human breast, particularly those diseases with polygenic  
CC aetiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
CC breast disease and non-carcinoma tumours. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

## SQ Sequence 50 AA;

Query Match 100.0%; Score 189; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1e-18;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36  
|||||

Db 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36  
|||||

## RESULT 12

ABG36903  
ID ABG36903 standard; peptide; 50 AA.  
XX  
AC ABG36903;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human peptide encoded by genome-derived single exon probe SEQ ID 26568.  
XX  
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemoderiosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease.  
XX  
OS Homo sapiens.  
XX  
PN W0200186003-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000665.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2002-114183/15.  
XX  
PT Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples.  
XX  
PS Claim 27; SEQ ID NO 26568; 634pp; English.  
XX

CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of probes  
CC; the novel set of probes which hybridise at high stringency to a nucleic  
CC acid expressed in the human lung; measuring gene expression in a sample  
CC derived from human lung, comprising (a) contacting the array with a  
CC collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of the  
CC array; identifying exons in a eukaryotic genome, comprising (a)  
CC algorithmically predicting at least one exon from genomic sequences of  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene expression  
CC analysis, and for identifying exons in a gene, particularly using human  
CC lung derived mRNA and for the study of lung diseases such as asthma, lung



```

XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX Claim 15; SEQ ID NO 21294; 530pp; English.
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 50 AA;
XX Query Match 100.0%; Score 189; DB 4; Length 50;
XX Best Local Similarity 100.0%; Pred. NO. 1e-18;
XX Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
RESULT 7
ID ABB19524 standard; protein; 50 AA.
XX ABB19524;
XX 23-JAN-2002 (first entry)
XX Protein #1523 encoded by probe for measuring heart cell gene expression.
XX Human; gene expression; heart; microarray; vascular system;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
XX Homo sapiens.
XX WO200157274-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX Claim 27; SEQ ID NO 11881; 327pp + Sequence Listing; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 50 AA;
XX Query Match 100.0%; Score 189; DB 4; Length 50;
XX Best Local Similarity 100.0%; Pred. NO. 1e-18;
XX Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
RESULT 8
ID AAM67252 standard; protein; 50 AA.
XX AAM67252;
XX 06-NOV-2001 (first entry)
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 27558.
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX Homo sapiens.
XX WO200157276-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA

```

```

XX AC AAM27545;
XX DT 17-OCT-2001 (first entry)
XX DE Peptide #1592 encoded by probe for measuring placental gene expression.
XX DE DE
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
XX PS Claim 27; SEQ ID NO 27814; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP:
XX CC see AAL31315-AA157546). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for producing a microarray for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from human placenta. The probes are useful for antenatal diagnosis of
XX CC human genetic disorders
XX SQ Sequence 50 AA;
Query Match 100.0%; Score 189; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
RESULT 5
ABB32389
ID ABB32389 standard; peptide; 50 AA.
XX AC ABB32389;
XX DT 01-FEB-2002 (first entry)
XX DE Peptide #5040 encoded by breast cell single exon nucleic acid probe.
XX DE DE
XX KW Human; microarray; single exon probe; gene expression; breast; disease;
XX KW cancer.
XX OS Homo sapiens.
XX PN WO200157271-A2.
XX PD 09-AUG-2001.

```

```

XX PF 30-JAN-2001; 2001WO-US000662.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX PT New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX PS Claim 27; SEQ ID NO 15357; 327pp + Sequence Listing; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each
XX CC probe of the microarray. The probes are useful for verifying the
XX CC expression of regions of genomic DNA predicted to encode proteins. They
XX CC are useful for gene discovery, and for determining predisposition and/or
XX CC prognosing breast disease. Gene expression analysis is useful for
XX CC assessing the toxicity of chemical agents on cells. The microarray of
XX CC this invention presents a far greater diversity of probes for measuring
XX CC gene expression, with far less bias than expressed sequence tag
XX CC microarrays. The method is suitable for rapid production of functional
XX CC information from genomic sequence. The present sequence is a peptide
XX CC encoded by a single exon nucleic acid probe of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 50 AA;
Query Match 100.0%; Score 189; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
RESULT 6
ABB28913
ID ABB28913 standard; peptide; 50 AA.
XX AC ABB28913;
XX DT 01-FEB-2002 (first entry)
XX DE Peptide #1564 encoded by breast cell single exon nucleic acid probe.
XX DE DE
XX KW Human; microarray; single exon probe; gene expression; breast; disease;
XX KW cancer.
XX OS Homo sapiens.
XX PN WO200157271-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000662.

```

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 XX  
 XX

SQ Sequence 36 AA;

Query Match 100.0%; Score 189; DB 5; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-19;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36  
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36

RESULT 2

AA015096  
 ID AAM15096 standard; protein; 50 AA.  
 AC AAM15096;  
 XX  
 XX  
 DT 12-OCT-2001 (first entry)

DE Peptide #1530 encoded by probe for measuring cervical gene expression.

XX Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 PN WO200157278-A2.

09-AUG-2001

PF 30-JAN-2001; 2001WO-US000670.

XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human cervical epithelial cells.  
 XX

PS Claim 27; SEQ ID NO 19922; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes  
 CC (SENPs; see AAL10068-AA128459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 50 AA;

Query Match 100.0%; Score 189; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1e-18;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36  
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36

RESULT 3

ABB34086  
 ID ABB34086 standard; peptide; 50 AA.  
 XX  
 AC ABB34086;  
 XX

DT 04-FEB-2002 (first entry)

XX Peptide #1592 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human foetal liver.  
 XX

PS Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring  
 CC human gene expression in a sample derived from human foetal liver. The  
 CC single exon nucleic acid probes may be used for predicting, measuring and  
 CC displaying gene expression in samples derived from human foetal liver. The  
 CC present sequence is a peptide encoded by a single exon nucleic acid probe  
 CC of the invention. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 50 AA;

Query Match 100.0%; Score 189; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1e-18;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36  
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36

RESULT 4

AA027545  
 ID AAM27545 standard; protein; 50 AA.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 32.6429 Seconds  
(without alignments)  
311.606 Million cell updates/sec

Title: US-09-941-314-6  
Perfect score: 189  
Sequence: 1 RKKTFLSVHEVAVENYAKDSLQWITDQYNKESDDK 36

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Genesep\_29Jan04:.\*  
1: Genesep1980s:.\*  
2: Genesep1990s:.\*  
3: Genesep2000s:.\*  
4: Genesep2001s:.\*  
5: Genesep2002s:.\*  
6: Genesep2003as:.\*  
7: Genesep2003bs:.\*  
8: Genesep2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	36	5 AAU79856	Aau79856 Human cys
2	189	100.0	50	4 AAM15096	Aam15096 Peptide #
3	189	100.0	50	4 ABB34086	Abb34086 Peptide #
4	189	100.0	50	4 AAM27545	Aam27545 Peptide #
5	189	100.0	50	4 ABB32389	Abb32389 Peptide #
6	189	100.0	50	4 ABB28913	Abb28913 Peptide #
7	189	100.0	50	4 ABB19524	Abb19524 Protein #
8	189	100.0	50	4 AAM67252	Aam67252 Human bon
9	189	100.0	50	4 AAM54871	Aam54871 Human bra
10	189	100.0	50	4 ABB48915	Abb48915 Human liv
11	189	100.0	50	4 AAM02833	Aam02833 Peptide #
12	189	100.0	50	5 ABB36903	Abb36903 Human pep
13	189	100.0	115	5 AAU79853	Aau79853 Human cys
14	189	100.0	117	5 AAU79854	Aau79854 Human cys
15	189	100.0	137	5 AAU79852	Aau79852 Human cys
16	101	53.4	22	5 AAU79855	Aau79855 Human cys
17	99	52.4	18	5 AAU79857	Aau79857 Human cys
18	99	52.4	35	5 AAU79858	Aau79858 Human cys
19	59.5	31.5	430	6 ABB30233	Abb30233 Protein e
20	57.5	30.4	184	5 ABB90713	Abb90713 Chlamydia
21	57	30.2	623	7 ADC95099	Adc95099 B. faeciu
22	56.5	29.9	97	6 ABBM73469	Abm73469 Staphyloc
23	56.5	29.9	579	6 ABBM16559	Abm16559 Protein e
24	55	29.1	27	5 AAU79859	Aau79859 Human cys
25	55	29.1	49	5 AAU79863	Aau79863 Human cys

# ALIGNMENTS

## RESULT 1

AAU79856  
ID AAU79856 standard; peptide; 36 AA.  
XX  
AC AAU79856;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human cystatin-8 (Zcys8) antigenic fragment #4.  
XX  
KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic fragment.  
XX  
OS Homo sapiens.  
XX  
PN WO200220567-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US026868.  
XX  
PR 01-SEP-2000; 2000US-0230230P.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Holloway JL, Gao Z, Bishop PD;  
DR WPI; 2002-383044/41.  
XX  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX  
PS Claim 2; Page 95; 100pp; English.  
XX  
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (i) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(i) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large

Abb09526 Soybean f  
Abu44069 Protein e  
Aay35198 Chlamydia  
Abu28659 Protein e  
Aar10545 Recombina  
Abu44692 Protein e  
Adc01083 Enterohae  
Abu47755 Protein e  
Aag82725 S. epider  
Abp39544 Staphyloc  
Aay00873 Original  
Aay00870 S. tubero  
Abb57153 Mouse isc  
Abu36321 Protein e  
Abu50049 Protein e  
Aag22389 Arabidops  
Aag51291 Arabidops  
Ada36895 Acinetoba  
Aag51290 Arabidops  
Aag22388 Arabidops

26 55 29.1 256 5 ABB09526  
27 54 28.6 239 6 ABU44069  
28 53.5 28.3 195 2 AAY35198  
29 53 28.0 240 6 ABU28659  
30 52 27.5 918 2 AAR10545  
31 51.5 27.2 278 6 ABU44692  
32 51.5 27.2 318 7 ADC01083  
33 51.5 27.2 418 6 ABU47755  
34 51 27.0 192 4 AAG82725  
35 51 27.0 277 5 ABP39544  
36 51 27.0 458 2 AAY00873  
37 51 27.0 766 2 AAY00870  
38 51 27.0 838 5 ABB57153  
39 51 27.0 1818 6 ABU36321  
40 50.5 26.7 460 6 ABU50049  
41 50 26.5 391 3 AAG22389  
42 50 26.5 391 3 AAG51291  
43 50 26.5 404 6 ADA36895  
44 50 26.5 409 3 AAG51290  
45 50 26.5 409 3 AAG22388

---

**THIS PAGE BLANK (USPTO)**

```

RESULT 13
US-10-424-599-249999
; Sequence 249999, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 249999
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT3847 67779C.1.dep

```

Search completed: March 18, 2004, 14:30:53  
Job time : 22.7857 secs

Db 1 KOSLOWITDOYNKESDDK 18

## RESULT 9

US-09-941-314-8  
; Sequence 8, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-8

Query Match 52.4%; Score 99; DB 9; Length 35;  
Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KOSLOWITDOYNKESDDK 36

Db 1 KOSLOWITDOYNKESDDK 18

## RESULT 10

US-10-282-122A-58157  
; Sequence 58157, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 58157  
; LENGTH: 430  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
US-10-282-122A-58157

Query Match 31.5%; Score 59.5; DB 12; Length 430;  
Best Local Similarity 52.4%; Pred. No. 9.5;  
Matches 11; Conservative 6; Mismatches 3; Indels 1; Gaps 1;

QY 14 VENZA-KOSLOWITDOYNKES 33

Db 169 LENYAKKEKLNWITDESNDN 189

## RESULT 11

US-10-312-273-375  
; Sequence 375, Application US/10312273  
; Publication No. US20040005667A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE  
; FILE REFERENCE: P025035WO  
; CURRENT APPLICATION NUMBER: US/10/312,273  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: 0016363.4  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: 0017047.2  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 0017983.8  
; PRIOR FILING DATE: 2000-07-21  
; PRIOR APPLICATION NUMBER: 0019368.0  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: 0020440.4  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: 0022583.9  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 0027549.5  
; PRIOR FILING DATE: 2000-11-10  
; PRIOR APPLICATION NUMBER: 0031706.5  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 664  
; SOFTWARE: SeqWin99, version 1.02  
; SEQ ID NO 375  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-10-312-273-375

Query Match 30.4%; Score 57.5; DB 15; Length 184;  
Best Local Similarity 38.5%; Pred. No. 7;  
Matches 15; Conservative 6; Mismatches 11; Indels 7; Gaps 1;

QY 1 RKKTFLSVHEVMAVENYAKD-----SLQWITDOYNKE 32

Db 26 RKKTFLSSHRVLARPSFPVDCPGKIYDLQEIYEELNAQ 64

## RESULT 12

US-10-282-122A-44483  
; Sequence 44483, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith

```

US-09-941-314-3
; Sequence 3, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-3

Query Match      100.0%; Score 189; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 6.5e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
      |||||||||||||||||||||||||||||||||||
Db      4 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 39

RESULT 5
US-09-941-314-4
; Sequence 4, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-4

Query Match      100.0%; Score 189; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 6.7e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
      |||||||||||||||||||||||||||||||||||
Db      6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 41

RESULT 6
US-09-941-314-2
; Sequence 2, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-2

Query Match      100.0%; Score 189; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 7.9e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
      |||||||||||||||||||||||||||||||||||
Db      26 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 61

RESULT 7
US-09-941-314-5
; Sequence 5, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-5

Query Match      53.4%; Score 101; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RKKTFLSVHEVMAVENYAKD 20
      |||||||||||||||
Db      3 RKKTFLSVHEVMAVENYAKD 22

RESULT 8
US-09-941-314-7
; Sequence 7, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-7

Query Match      52.4%; Score 99; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 KDSLQWITDQYNKESDDK 36
      |||||||||||

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 21.7857 Seconds  
(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-6

Perfect score: 189

Sequence: 1 RKKTFSLVHEVAVENYAKDSLQWITDQYNKESDDK 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	36	9 US-09-941-314-6	Sequence 6, Appli
2	189	100.0	50	9 US-09-864-761-34822	Sequence 34822, A
3	189	100.0	50	9 US-09-864-761-48936	Sequence 48936, A
4	189	100.0	115	9 US-09-941-314-3	Sequence 3, Appli
5	189	100.0	117	9 US-09-941-314-4	Sequence 4, Appli
6	189	100.0	137	9 US-09-941-314-2	Sequence 2, Appli
7	101	53.4	22	9 US-09-941-314-5	Sequence 5, Appli
8	99	52.4	18	9 US-09-941-314-7	Sequence 7, Appli
9	99	52.4	35	9 US-09-941-314-8	Sequence 8, Appli
10	59.5	31.5	430	12 US-10-282-122A-58157	Sequence 58157, A
11	57.5	30.4	184	15 US-10-312-273-375	Sequence 375, App
12	56.5	29.9	579	12 US-10-282-122A-44483	Sequence 44483, A
13	55.5	29.4	131	12 US-10-424-599-249999	Sequence 249999,
14	55	29.1	27	9 US-09-941-314-9	Sequence 9, Appli
15	55	29.1	49	9 US-09-941-314-13	Sequence 13, Appli

16	55	29.1	252	12	US-10-424-599-261063	Sequence 261063,
17	55	29.1	260	12	US-10-424-599-180261	Sequence 180261,
18	54.5	28.8	186	12	US-10-424-599-275467	Sequence 275467,
19	54.5	28.8	369	12	US-10-425-114-56041	Sequence 56041, A
20	54.5	28.8	486	12	US-10-424-599-275468	Sequence 275468,
21	54.5	28.8	507	12	US-10-425-114-57763	Sequence 57763, A
22	54	28.6	239	12	US-10-282-122A-71993	Sequence 71993, A
23	54	28.6	684	12	US-10-425-114-46564	Sequence 46564, A
24	53.5	28.3	195	15	US-10-289-762-616	Sequence 616, App
25	53.5	28.3	489	12	US-10-424-599-275473	Sequence 275473,
26	53.5	28.3	494	12	US-10-425-114-45919	Sequence 45919, A
27	53	28.0	240	12	US-10-282-122A-56583	Sequence 56583, A
28	51.5	27.2	278	12	US-10-282-122A-72616	Sequence 72616, A
29	51.5	27.2	282	12	US-10-425-114-64859	Sequence 64859, A
30	51.5	27.2	384	12	US-10-425-114-68692	Sequence 68692, A
31	51.5	27.2	418	12	US-10-282-122A-75679	Sequence 75679, A
32	51	27.0	81	12	US-10-425-114-72398	Sequence 72398, A
33	51	27.0	159	12	US-10-424-599-259133	Sequence 259133,
34	51	27.0	669	12	US-10-424-599-178326	Sequence 178326,
35	51	27.0	1818	12	US-10-282-122A-64245	Sequence 64245, A
36	50.5	26.7	460	12	US-10-282-122A-77973	Sequence 77973, A
37	50	26.5	430	15	US-10-320-800-62	Sequence 62, Appli
38	50	26.5	469	12	US-10-282-122A-50269	Sequence 50269, A
39	50	26.5	512	12	US-10-425-114-45567	Sequence 45567, A
40	50	26.5	626	12	US-10-425-114-58446	Sequence 58446, A
41	50	26.5	669	12	US-10-424-599-178328	Sequence 178328,
42	50	26.5	781	12	US-10-424-599-178322	Sequence 178322,
43	50	26.5	1166	15	US-10-104-047-2949	Sequence 2949, Ap
44	50	26.5	1175	12	US-10-282-122A-57562	Sequence 57562, A
45	50	26.5	1208	9	US-09-815-242-10627	Sequence 10627, A

#### ALIGNMENTS

RESULT 1  
US-09-941-314-6  
; Sequence 6, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-6

Query Match 100.0%; Score 189; DB 9; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKKTFSLVHEVAVENYAKDSLQWITDQYNKESDDK 36  
Db 1 RKKTFSLVHEVAVENYAKDSLQWITDQYNKESDDK 36

RESULT 2  
US-09-864-761-34822  
; Sequence 34822, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.



Query Match 28.6%; Score 54; DB.2; Length 663;

K/III-Worram, A.; Sutteler, C.; Ozersky, E. submitted to the EMBL Data Library, May 1999

Best Local Similarity 34.5%; Pred. No. 2.9; Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 3 KTFLSVHEVMAVENYAKDSLQWITDQYNK 31  
DB 74 KNFLAIDPVLNPFNSQHLMWNDLDFNE 102

RESULT 3

G85808  
hypothetical protein sdia [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C;Accession: G85808  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: G85808  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-240 <STO>  
A;Cross-references: GB:AE005174; NID:gl21516014; PIDN:AAG56931.1; GSPDB:GN001145; UWGP:Z30  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: sdia  
C;Superfamily: sdia regulatory protein

Query Match 30.7%; Score 58; DB 2; Length 240;  
Best Local Similarity 34.5%; Pred. No. 2.9; Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 3 KTFLSVHEVMAVENYAKDSLQWITDQYNK 31  
DB 74 KNFLAIDPVLNPFNSQHLMWNDLDFNE 102

RESULT 4

D72061  
hypothetical protein CP0175 [imported] - Chlamydomophila pneumoniae (strains CWL029 and AR  
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
C;Accession: D72061; A81606  
R;Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999  
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A;Reference number: A72000; MUID:99206606; PMID:10192388  
A;Accession: D72061  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-184 <ARN>  
A;Cross-references: GB:AE001642; GB:AE001363; NID:g4376865; PIDN:AAD18714.1; PID:g4376865  
A;Experimental source: strain CWL029  
R;Read, T.D.; Brunnham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A;Reference number: A81500; MUID:20150255; PMID:10684935  
A;Accession: A81606  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-184 <REA>  
A;Cross-references: GB:AE002178; GB:AE002161; NID:g7189099; PIDN:AAF38049.1; PID:g718910  
A;Experimental source: strain AR39, HL cells  
C;Genetics:  
A;Gene: CPn0574; CP0175

Query Match 30.4%; Score 57.5; DB 2; Length 184;  
Best Local Similarity 38.5%; Pred. No. 2.5; Matches 15; Conservative 6; Mismatches 11; Indels 7; Gaps 1;

QY 1 RKKTFLSVHEVMAVENYAKD-----SLOWITDQYNKE 32

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 6.92857 Seconds  
(without alignments)  
499.799 Million cell updates/sec

Title: US-09-941-314-6  
Perfect score: 189  
Sequence: 1 RKKTFLSVHEVMAVENYAKSLQWITDQYNKESDDX 36

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: Piri:\*  
2: Piri2:\*  
3: Piri3:\*  
4: Piri4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.5	31.5	430	2 D64151	hypothetical prote
2	58	30.7	240	2 F90960	hypothetical prote
3	58	30.7	240	2 G85808	hypothetical prote
4	57.5	30.4	184	2 D72061	hypothetical prote
5	57.5	30.4	184	2 B86562	hypothetical prote
6	56	29.6	209	2 T20975	hypothetical prote
7	55	29.1	343	2 AG2856	hypothetical prote
8	54	28.6	467	2 JQ0966	dnak-type molecula
9	54	28.6	663	2 T04078	dnak-type molecula
10	54	28.6	663	2 T04080	dnak-type molecula
11	53	28.0	195	2 S28845	myosin regulatory
12	53	28.0	205	2 T33117	hypothetical prote
13	53	28.0	240	1 Q0SCU1	probable transcrip
14	53	28.0	403	1 Q9SA4E	hypothetical prote
15	52	27.5	404	2 C96640	hypothetical prote
16	52	27.5	457	2 T05439	hypothetical prote
17	52	27.5	1422	2 T18404	chromatin remodel
18	51.5	27.2	318	2 C85517	unknown protein en
19	51.5	27.2	318	2 H90666	hypothetical prote
20	51.5	27.2	418	1 BVEBHA	glutamy1-trNA redu
21	51.5	27.2	418	2 AC0720	glutamy1-trNA redu
22	51.5	27.2	465	2 A81419	hypothetical prote
23	51.5	27.2	567	2 B69166	hypothetical prote
24	51.5	27.2	656	2 A48439	dnak-type molecula
25	51	27.0	190	2 F97008	hypothetical prote
26	51	27.0	290	2 S21877	dnak-type molecula
27	51	27.0	292	2 S21878	dnak-type molecula
28	51	27.0	293	2 S21881	dnak-type molecula
29	51	27.0	326	2 T45226	probable N5,N10-me

30	51	27.0	360	2 C72356	hypothetical prote
31	51	27.0	414	2 B81413	probable two-compo
32	51	27.0	420	2 S67559	probable membrane
33	51	27.0	438	2 B71021	probable L-asparag
34	51	27.0	666	2 T06358	dnak-type molecula
35	51	27.0	667	2 S21879	dnak-type molecula
36	51	27.0	668	2 S21880	dnak-type molecula
37	51	27.0	668	2 T46574	hypothetical prote
38	51	27.0	1818	1 S73852	conserved hypotet
39	50.5	26.7	460	2 AE0130	hypothetical prote
40	50.5	26.7	732	2 T32757	conserved hypotet
41	50	26.5	252	2 H71292	hypothetical prote
42	50	26.5	302	2 T31542	hypothetical prote
43	50	26.5	430	2 A81790	conserved hypotet
44	50	26.5	430	2 A81214	conserved hypotet
45	50	26.5	525	2 E96786	protein F10A5.13 (

ALIGNMENTS

RESULT 1

D64151  
hypothetical protein HI0404 - Haemophilus influenzae (strain Rd KW20)  
C;Species: Haemophilus influenzae  
C;Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Dec-2002  
C;Accession: D64151

R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995

A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A;Reference number: A64000; MUID:95350630; PMID:7542800  
A;Accession: D64151

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-430 <TIGR>

A;Cross-references: GB:U32723; GB:L42023; NID:gl573363; PIDN:AAC22063.1; PID:gl573375; TGI

A;Note: best homolog was a hypothetical protein from Bacillus subtilis

C;Superfamily: Cell cycle protein Mesaj

Query Match 31.5%; Score 59.5; DB 2; Length 430;  
Best Local Similarity 52.4%; Pred. No. 3.6;  
Matches 11; Conservative 6; Mismatches 3; Indels 1; Gaps 1;

Qy 14 VENYA-KDSLQWITDQYNKES 33

||||| : ||||| : |||

Db 169 LENYAQKEKLWITDSSNEDN 189

RESULT 2

F90960  
hypothetical protein ECe2654 [imported] - Escherichia coli (strain O157:H7, substrain RIN

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001

C;Accession: F90960

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F90960

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-240 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA036077.1; PID:gl13362122; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECe2654

C;Superfamily: sdiA regulatory protein

Query Match 30.7%; Score 58; DB 2; Length 240;

Search completed: March 18, 2004, 14:25:35  
Job time : 9.21429 secs

**THIS PAGE BLANK (cont.)**

Query Match 25.9%; Score 49; DB 2; Length 488;  
Best Local Similarity 31.2%; Pred. No. 56;  
Matches 10; Conservative 8; Mismatches 14; Indels

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6509155 g312212  
US-09-507-765-33

Query Match 26.2%; Score 49.5; DB 4; Length 439;  
Best Local Similarity 39.4%; Pred. No. 41;  
Matches 13; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY 4 TELSVEHMAVENYAKDSLQWITDQYNKESDDK 36  
DB 121 TLLYHLHGLTSDN--KPSUSWLRDAY-REFDRK 150

## RESULT 10

US-09-198-452A-638  
; Sequence 638, Application US/09198452A

; Patent No. 6559294

; GENERAL INFORMATION:

; APPLICANT: Griffais, R.

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention

; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection

; FILE REFERENCE: 9710-003-999

; CURRENT APPLICATION NUMBER: US/09/198,452A

; CURRENT FILING DATE: 1998-11-24

; NUMBER OF SEQ ID NOS: 6849

; SEQ ID NO 638

; LENGTH: 460

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-198-452A-638

Query Match 26.2%; Score 49.5; DB 4; Length 460;  
Best Local Similarity 35.9%; Pred. No. 44;  
Matches 14; Conservative 10; Mismatches 8; Indels 7; Gaps 3;

QY 2 KKTFLSVHEVMAVENYAK--DSLQWITDQ---YNKESDD 35  
DB 172 KXSLLEIRQVL--EDYAKLSAISFTQDKLWIEKESD 208

## RESULT 11

US-09-392-714-24

; Sequence 24, Application US/09392714A

; Patent No. 6686147

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Matthew J.

; APPLICANT: Gure, Ali O.

; APPLICANT: Williamson, Barbara

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: Cancer Associated Antigens and Uses

; FILE REFERENCE: L0461/7062

; CURRENT APPLICATION NUMBER: US/09/392,714A

; CURRENT FILING DATE: 1999-09-09

; EARLIER APPLICATION NUMBER: PCT/US98/14679

; EARLIER FILING DATE: 1998-07-15

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 24

; LENGTH: 973

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-392-714-24

Query Match 26.2%; Score 49.5; DB 4; Length 973;  
Best Local Similarity 37.8%; Pred. No. 1.1e+02;  
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 6 LSVHEV-----MAVENYAKDSLQWITDQYNKESDDK 36

DB 633 LNVYEIKVNKLELES-AKQKFGTDTYQKEIEDK 668

## RESULT 12

US-09-104-324B-4

; Sequence 4, Application US/09104324B

; Patent No. 6232460

; GENERAL INFORMATION:

; APPLICANT: T recti, Ozlem; Sahin, Ugur; Pfreundschuh, Michael

; TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers,

; TITLE OF INVENTION: And Methods For Identifying Pathogenic Markers In A Sample Of

; TITLE OF INVENTION: No. 6232460mal Cells

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fulbright & Jaworski LLP

; STREET: 666 Fifth Avenue

; CITY: New York City

; STATE: New York

; ZIP: 10103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/104,324B

; FILING DATE: 25-June-1998

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/892,702

; FILING DATE: 15-July-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 6232460man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5491

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 318-3000

; TELEFAX: (212) 752-5958

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 976 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-09-104-324B-4

Query Match 26.2%; Score 49.5; DB 3; Length 976;  
Best Local Similarity 37.8%; Pred. No. 1.1e+02;  
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 6 LSVHEV-----MAVENYAKDSLQWITDQYNKESDDK 36

DB 636 LNVYEIKVNKLELES-AKQKFGTDTYQKEIEDK 671

## RESULT 13

US-08-913-578-2

; Sequence 2, Application US/08913578

; Patent No. 6218159

; GENERAL INFORMATION:

; APPLICANT: Hodgson, John

; APPLICANT: Lawlor, Elizabeth

; TITLE OF INVENTION: No. 6218159el tRNA synthetase

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406-0939

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

Query Match	26.5%;	Score 50;	DB 4;	Length 1224;
Best Local Similarity	35.3%;	Pred. No. 1.2e+02;		

```

1  RESULT 9
2  US-09/507-765-33
3  ; Sequence 33, Application US/09507765
4  ; Patent No. 6509155
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Klinger, Tod M.
7  ; APPLICANT: Stewart, Elizabeth A.
8  ; APPLICANT: Yue, Henry
9  ; APPLICANT: Baughn, Mariah R.
10 ; TITLE OF INVENTION: GTPASE ACTIVATING PROTEINS
11 ; FILE REFERENCE: PC-0010 US
12 ; CURRENT APPLICATION NUMBER: US/09/507,765
13 ; CURRENT FILING DATE: 2000-02-18
14 ; NUMBER OF SEQ ID NOS: 33
15 ; SOFTWARE: PERL Program
16 ; SEQ ID NO 33
17 ; LENGTH: 439
18 ; TYPE: PRT

```

```
Query Match      30.2%; Score 57; DB 4; Length 623;
Best Local Similarity 35.5%; Pred. No. 5.1;
Matches 11; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY      6 LSHEVMAVENYAKDSLOWITDQYNKESDDK 36
DB      225 LKVEPMDEPGYSGEKIKETIKYQKLADSK 255

RESULT 2
US-09-198-452A-616
; Sequence 616, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 616
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-616

Query Match      28.3%; Score 53.5; DB 4; Length 195;
Best Local Similarity 35.9%; Pred. No. 3.9;
Matches 14; Conservative 7; Mismatches 11; Indels 7; Gaps 1;

QY      1 RKKTFLSVHEVMAVENYAKD-----SLQWITDQYNKE 32
DB      37 RQKTFLSHSHVLRARPFVDYCPGKIYDLQEIYEELNAQ 75

RESULT 3
US-09-134-001C-4389
; Sequence 4389, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4389
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4389

Query Match      27.0%; Score 51; DB 4; Length 277;
Best Local Similarity 38.7%; Pred. No. 14;
Matches 12; Conservative 6; Mismatches 7; Indels 6; Gaps 1;

QY      1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNK 31
DB      204 RKKPFSLMPPVAMVYSL-----WLTKKKYKK 228

RESULT 4
US-09-463-238-15
; Sequence 15, Application US/09463238
; Patent No. 6469230
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
```

```
; GENERAL INFORMATION:
; APPLICANT: Edwards, Elizabeth A
; APPLICANT: Smith, Alison M
; APPLICANT: Bustos Guillen, Regla
; APPLICANT: Martin, Catherine R
; APPLICANT: Plant Bioscience Limited
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 97.118
; CURRENT APPLICATION NUMBER: US/09/463,238
; CURRENT FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/GB98/02280
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: GB 9716185.5
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-463-238-15

Query Match      27.0%; Score 51; DB 4; Length 397;
Best Local Similarity 40.7%; Pred. No. 22;
Matches 11; Conservative 6; Mismatches 8; Indels 2; Gaps 2;

QY      9 HEVMAVENY-AKDSLOWITDQ-YNKES 33
DB      292 HNVLRKENFIEKNDITWLEDNWNYES 318

RESULT 5
US-09-463-238-6
; Sequence 6, Application US/09463238
; Patent No. 6469230
; GENERAL INFORMATION:
; APPLICANT: Edwards, Elizabeth A
; APPLICANT: Smith, Alison M
; APPLICANT: Bustos Guillen, Regla
; APPLICANT: Martin, Catherine R
; APPLICANT: Plant Bioscience Limited
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 97.118
; CURRENT APPLICATION NUMBER: US/09/463,238
; CURRENT FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/GB98/02280
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: GB 9716185.5
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-463-238-6

Query Match      27.0%; Score 51; DB 4; Length 766;
Best Local Similarity 40.7%; Pred. No. 50;
Matches 11; Conservative 6; Mismatches 8; Indels 2; Gaps 2;

QY      9 HEVMAVENY-AKDSLOWITDQ-YNKES 33
DB      661 HNVLRKENFIEKNDITWLEDNWNYES 687

RESULT 6
US-09-328-352-8182
; Sequence 8182, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 9.21429 Seconds  
(without alignments)  
201.701 Million cell updates/sec

Title: US-09-941-314-6  
Perfect score: 189  
Sequence: 1 RKKTFSLVHEVAVENYAKDSLOWTDQYNKESDDK 36

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	30.2	623	4	US-09-107-532A-4726
2	53.5	28.3	195	4	US-09-198-452A-616
3	51	27.0	277	4	US-09-134-001C-4389
4	51	27.0	397	4	US-09-463-238-15
5	51	27.0	786	4	US-09-463-238-6
6	50	26.5	404	4	US-09-328-352-8182
7	50	26.5	1224	4	US-09-107-532A-6220
8	49.5	26.2	125	2	US-08-408-095-35
9	49.5	26.2	439	4	US-09-507-765-33
10	49.5	26.2	460	4	US-09-198-452A-638
11	49.5	26.2	973	4	US-09-392-714-24
12	49.5	26.2	976	3	US-09-104-324B-4
13	49	25.9	484	3	US-08-913-578-2
14	49	25.9	484	3	US-08-785-427-2
15	49	25.9	488	2	US-08-599-455B-5
16	49	25.9	488	3	US-09-069-781B-5
17	49	25.9	488	4	US-09-337-132-5
18	49	25.9	488	4	US-08-864-564A-5
19	49	25.9	488	4	US-09-094-410-5
20	49	25.9	488	4	US-08-708-123D-5
21	49	25.9	488	4	US-08-583-153A-5
22	49	25.9	488	4	US-08-638-524B-5
23	49	25.9	451	4	US-09-107-532A-6380
24	49	25.9	572	2	US-08-419-652-5
25	49	25.9	655	4	US-09-632-538C-36
26	49	25.9	658	2	US-08-825-558-4
27	49	25.9	658	4	US-09-312-611-4

28	49	25.9	708	1	US-07-797-556-2	Sequence 2, Appli
29	49	25.9	708	1	US-08-308-881-2	Sequence 2, Appli
30	49	25.9	708	2	US-09-058-263-2	Sequence 2, Appli
31	49	25.9	708	2	US-09-059-099-2	Sequence 2, Appli
32	49	25.9	708	3	US-09-058-264-2	Sequence 2, Appli
33	49	25.9	708	4	US-09-455-962-2	Sequence 2, Appli
34	49	25.9	708	5	PCT-US95-06530-2	Sequence 2, Appli
35	49	25.9	859	4	US-09-313-942-7	Sequence 7, Appli
36	49	25.9	918	2	US-08-825-558-6	Sequence 6, Appli
37	49	25.9	918	4	US-09-312-611-6	Sequence 6, Appli
38	49	25.9	951	4	US-09-313-942-9	Sequence 9, Appli
39	49	25.9	1002	4	US-09-762-724-4	Sequence 4, Appli
40	49	25.9	1158	4	US-09-313-942-26	Sequence 26, Appli
41	49	25.9	1168	4	US-09-313-942-24	Sequence 24, Appli
42	48.5	25.7	1128	1	US-07-882-790-4	Sequence 4, Appli
43	48.5	25.7	1129	4	US-09-543-681A-6416	Sequence 6416, Ap
44	48	25.4	371	4	US-09-653-375B-9	Sequence 9, Appli
45	48	25.4	803	4	US-09-540-236-2588	Sequence 2588, Ap

ALIGNMENTS

RESULT 1

US-09-107-532A-4726  
; Sequence 4726, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 4726:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 623 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (B) LOCATION 1...623  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4726:

US-09-107-532A-4726

DE ORF27.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=569b;  
RX MEDLINE=20476433; PubMed=11021935;  
RA Clark C.A., Putins L., Kaewrakon P., Focareta T., Manning P.A.;  
RT "The Vibrio cholerae O1 chromosomal integron."  
RL Microbiology 146:2605-2612(2000).  
DR EMBL; AFI79596; AAF71189.1; -.  
SQ SEQUENCE 301 AA; 33413 MW; F0D66AF91E11609F CRC64;

Query Match 40.0%; Score 44; DB 2; Length 301;  
Best Local Similarity 47.4%; Pred.No. 48;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 ARKKTFLSVHEVMAVENYA 20  
|||:|:|:|:  
Db 208 ARVNGFSSIHEMNGTEVFA 226

Search completed: March 18, 2004, 14:21:25  
Job time : 15.1825 secs

THIS PAGE BLANK (USPIC)

[illegible]

```

DR InterPro; IPR002550; CBS.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF01595; DUF21; 1.
SQ SEQUENCE 706 AA; 77711 MW; AAB3EACAD8E5F53 CRC64;

Query Match      40.9%; Score 45; DB 3; Length 706;
Best Local Similarity 44.4%; Pred. No. 80;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      5 KTELVSHEVMAVENYAKD 22
Db      222 KTLVTLHRTMGVERLTKD 239

RESULT 9
QNG54      PRELIMINARY; PRT; 852 AA.
AC QNG54;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aeryl hydrocarbon receptor-like protein.
OS Mya arenaria.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Myoidea; Myidae; Mya.
OX NCBI_TaxID=5604;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley M.L., Powell W.H., Hahn M.E., Van Beneden R.J.;
RT "An Aeryl Hydrocarbon Receptor from the Softshell Clam, Mya arenaria.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
    TRANSCRIPTION FACTORS.
DR EMBL; AF261769; AAF70378.1; -
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001610; PAC_basic.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 1.
DR PROSITE; PS00888; HLH_2; 1.
DR PROSITE; PS00112; PAS; 1.
KW Receptor.
SQ SEQUENCE 852 AA; 96580 MW; 06353220C22ABC10 CRC64;

Query Match      40.9%; Score 45; DB 5; Length 852;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      5 KTELVSHEVMAVENY 19
Db      186 KANLTLHEVMPENY 200

RESULT 10
Q8P213      PRELIMINARY; PRT; 446 AA.
AC Q8P213;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative DNA polymerase III delta prime subunit (Putative DNA
    polymerase III delta prime subunit-phage associated).
GN SPY18_0348 OR SPY18_0690.
OS Streptococcus pyogenes (serotype M18), and
    Streptococcus pyogenes (serotype M3).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103, 198466;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
    Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
    Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
    Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
    group A Streptococcus strains associated with acute rheumatic fever
    outbreaks.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=1212206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
    Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
    Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
    Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
    phage-encoded toxins, the high-virulence phenotype, and clone
    emergence.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [3]
DR EMBL; AS009979; AAL97100.1; -
DR EMBL; AE014150; AAM79297.1; -
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003678; F:DNA helicase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR007694; DnaB_C.
DR InterPro; IPR007692; DnaB_helicase.
DR InterPro; IPR007693; DnaB_N.
DR Pfam; PF00772; DnaB; 1.
DR Pfam; PF03796; DnaB_C; 1.
DR ProDom; PD332834; DnaB_C; 1.
DR TIGRFAMs; TIGR00665; DnaB; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 446 AA; 50028 MW; FD98A785AAA3BE2E CRC64;

Query Match      40.5%; Score 44.5; DB 16; Length 446;
Best Local Similarity 45.0%; Pred. No. 60;
Matches 9; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY      2 ARKTEFLSVHEVMAVENYAK 21
Db      151 SNKSPFRPHDVLA-ENYOK 169

RESULT 11
Q878M1      PRELIMINARY; PRT; 451 AA.
AC Q878M1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative DNA polymerase III delta prime subunit.
GN SP51160.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
    Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
    Hayaashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
    of S. pyogenes SSI-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

```

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Chisoe S.;  
 RT "The sequence of C. elegans cosmid C33F10.";  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U49830; AAK31482.1; -;  
 DR WormPep; C33F10.5b; C25803.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR008957; FN\_III-like.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR Pfam; PF00041; fn3\_2.  
 DR Pfam; PF00047; ig; 2.  
 DR SMART; SM00060; FN3; 4.  
 DR SMART; SM00408; IGC2; 2.  
 DR PROSITE; PS0835; IG\_LIKE; 6.  
 KW Hypothetical protein; Immunoglobulin domain.  
 SQ SEQUENCE 1227 AA; 136318 MW; 1E808292F7BD2C3A CRC64;

Query Match 43.6%; Score 48; DB 5; Length 1227;  
 Best Local Similarity 57.9%; Pred. No. 44;  
 Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 KKTFLSVHVEVMAVENYAKD 22  
 ||||| : : : : :  
 DB 5 KKTILSAVEIDVGNKAAD 23

RESULT 6  
 Q8DB15 PRELIMINARY; PRT; 105 AA.  
 AC Q8DB15;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN V12013.  
 OS Vibrio vulnificus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=672;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CMCP6;  
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
 RA Choy H.E.;  
 RT "Complete genome sequence of Vibrio vulnificus CMCP6.";  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE016803; AA010409.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 105 AA; 12052 MW; 608171D5253EF918 CRC64;

Query Match 42.7%; Score 47; DB 16; Length 105;  
 Best Local Similarity 31.8%; Pred. No. 4.9;  
 Matches 7; Conservative 11; Mismatches 2; Indels 2; Gaps 1;

QY 1 QARKKFLSVHVEVMAVENYAKD 22  
 : : : : :  
 DB 61 ESKKSY--EHTILAKNFARD 80

RESULT 7  
 Q8Z245 PRELIMINARY; PRT; 273 AA.  
 AC Q8Z245;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Glycosyltransferase (Type 2).  
 GN PAE0441.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 OX NCBI\_TaxID=13773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX MEDLINE=21664397; PubMed=11792869;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 RT aerophilum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).  
 DR EMBL; AE009767; AAL62796.1; -;  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR InterPro; IPR001064; Crystallin.  
 DR InterPro; IPR001173; Glyco\_transf\_2.  
 DR Pfam; PF00535; Glycos\_transf\_2; 1.  
 DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
 KW Transferase; Complete proteome.  
 SQ SEQUENCE 273 AA; 31634 MW; A002488D5A68663B CRC64;

Query Match 41.4%; Score 45.5; DB 17; Length 273;  
 Best Local Similarity 45.0%; Pred. No. 24;  
 Matches 9; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

QY 3 RKKTFLSVHVEVMAVENYAKD 22  
 ||:|: : : : : : : : : :  
 DB 27 RKQTYKNI-EVIVVDNYSTD 45

RESULT 8  
 Q12296 PRELIMINARY; PRT; 706 AA.  
 AC Q12296;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Chromosome XV reading frame ORF YOL060C.  
 GN MAM3 OR AM13 OR YOL060C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Feldmann H., Mannhaupt G., Vetter I.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MTPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ALPHA S288;  
 RA Mannhaupt G., Vetter I., Schwarlose C., Mitzel S., Feldmann H.;  
 RL Yeast 0:0-0(0).  
 DR EMBL; Z74802; CAA99069.1; -;  
 DR EMBL; X91067; CAA62524.1; -;  
 DR PIR; S61717; S61717.  
 DR SGD; S0005421; MAM3.  
 DR GO; GO:0007005; P:mitochondrion organization and biogenesis; IMP.

```
DE SC13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335480; AAL71991.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 138 AA; 16506 MW; B49440ACA3585C64 CRC64;

Query Match 100.0%; Score 110; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QARKKTFLSVHEVMAVENYAKD 22
DB 25 QARKKTFLSVHEVMAVENYAKD 46

RESULT 3
O8K5A3 PRELIMINARY; PRT; 139 AA.
ID Q8K5A3
AC Q8K5A3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin 11.
GN CST11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RA Hamil K.G., Hall S.H.;
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF501290; AAM21709.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 139 AA; 16686 MW; B1E36DB786B4D08C CRC64;

Query Match 48.2%; Score 53; DB 11; Length 139;
Best Local Similarity 45.5%; Pred. No. 0.63;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 QARKKTFLSVHEVMAVENYAKD 22
DB 25 QVKKRTFIRVEVNALESSVYKE 76

RESULT 4
Q927X9 PRELIMINARY; PRT; 184 AA.
ID Q927X9
AC Q927X9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Cpn0574.
GN Cpn0574 OR CPJ0574 OR CP0175 OR CPB0596.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=9920606; PubMed=10192388;
RA Kalnan S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,
RA White O., Hickey E.K., Peterson J., Uettermann T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=TW-183;
RX Geng M.N., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS001642; AAD18714.1; -.
DR EMBL; AR002178; AAF38049.1; -.
DR EMBL; AF002547; BAA98780.1; -.
DR EMBL; AR017159; AAP98525.1; -.
DR PIR; B8562; B86562.
DR PIR; D72061; D72061.
DR TIGR; CP0175; -.
DR GO; GO:0008237; F:metalloproteinase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 184 AA; 22024 MW; EF36820D1A97A814 CRC64;

Query Match 48.2%; Score 53; DB 16; Length 184;
Best Local Similarity 55.0%; Pred. No. 0.84;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22
DB 26 RKKTFLSHRVLARPSFPVD 45

RESULT 5
Q9BIA2 PRELIMINARY; PRT; 1227 AA.
ID Q9BIA2
AC Q9BIA2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C33F10.5 OR C33F10.5B.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Nematoda; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 13.1825 Seconds

(without alignments)  
526.560 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKKTFSLVHEVMAVENYAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	103	4 Q8WXU5	Q8WXU5 homo sapien
2	110	100.0	138	4 Q8WXU6	Q8WXU6 homo sapien
3	53	48.2	139	11 Q8K5A3	Q8K5A3 rattus norv
4	53	48.2	184	16 Q9Z7X9	Q9Z7X9 chlamydia p
5	48	43.6	1227	5 Q9BIA2	Q9BIA2 caenorhabdi
6	47	42.7	105	16 Q8DB15	Q8DB15 vibrio vuln
7	45.5	41.4	273	17 Q8Z245	Q8Z245 pyrobaculum
8	45	40.9	706	3 Q12296	Q12296 saccharomyc
9	45	40.9	852	5 Q9NG54	Q9NG54 mya arenari
10	44.5	40.5	446	16 Q8P2I3	Q8P2I3 streptococc
11	44.5	40.5	451	16 Q878M1	Q878M1 streptococc
12	44.5	40.5	1025	5 Q8IC13	Q8IC13 plasmodium
13	44	40.0	175	16 Q81J78	Q81J78 bacillus an
14	44	40.0	176	16 Q814Q9	Q814Q9 bacillus ce
15	44	40.0	301	2 Q9L985	Q9L985 vibrio chol
16	44	40.0	542	4 Q96M62	Q96M62 homo sapien

17	44	40.0	576	4 Q81Y28	Q81Y28 homo sapien
18	44	40.0	704	4 Q96EF0	Q96EF0 homo sapien
19	43	39.1	102	10 Q49991	Q49991 lycopersico
20	43	39.1	127	9 Q8SCR2	Q8SCR2 pseudomonas
21	43	39.1	212	5 Q966N6	Q966N6 caenorhabdi
22	43	39.1	259	16 Q25194	Q25194 helicobacte
23	43	39.1	294	16 Q88BQ6	Q88BQ6 pseudomonas
24	43	39.1	366	5 Q93542	Q93542 caenorhabdi
25	43	39.1	394	2 Q8RK17	Q8RK17 pseudomonas
26	43	39.1	428	16 Q88C18	Q88C18 pseudomonas
27	43	39.1	431	16 Q9JXY5	Q9JXY5 neisseria m
28	43	39.1	431	16 Q9JYV6	Q9JYV6 neisseria m
29	43	39.1	467	5 Q19003	Q19003 caenorhabdi
30	43	39.1	488	10 Q9SEJ7	Q9SEJ7 lupinus alb
31	43	39.1	493	10 Q94E49	Q94E49 oryza sativ
32	43	39.1	585	16 Q97KY8	Q97KY8 clostridium
33	43	39.1	676	10 Q48785	Q48785 arabidopsis
34	43	39.1	1217	16 Q82241	Q82241 enterococcu
35	42	38.2	117	15 Q9YVW4	Q9YVW4 human immun
36	42	38.2	119	10 Q9S9D5	Q9S9D5 glycine max
37	42	38.2	122	15 Q90DQ6	Q90DQ6 human immun
38	42	38.2	127	15 Q9YZ16	Q9YZ16 human immun
39	42	38.2	140	5 Q8T5B0	Q8T5B0 plasmodium
40	42	38.2	154	16 Q8XI64	Q8XI64 clostridium
41	42	38.2	208	2 Q9F6B5	Q9F6B5 pseudoalter
42	42	38.2	209	2 Q68663	Q68663 neptunomona
43	42	38.2	254	10 Q84RU4	Q84RU4 oryza sativ
44	42	38.2	256	16 Q99RZ1	Q99RZ1 staphylococ
45	42	38.2	256	16 Q8NV94	Q8NV94 staphylococ

## ALIGNMENTS

### RESULT 1

Q8WXU5	PRELIMINARY;	PRT;	103 AA.
AC Q8WXU5	Q8WXU5		
DT 01-MAR-2002	(TREMBLrel. 20, Created)		
DT 01-MAR-2002	(TREMBLrel. 20, Last sequence update)		
DT 01-JUN-2003	(TREMBLrel. 24, Last annotation update)		
DE SC13delta.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;			
RT "SC13: A novel epididymal specific member of the cystatin family.";			
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AF335481; AAL71992.1; -			
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.			
DR InterPro; IPR00010; Cystatin.			
DR Pfam; PF00031; cystatin; 1.			
SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;			

Query Match 100.0%; Score 110; DB 4; Length 103;  
Best Local Similarity 100.0%; Pred. No. 8.8e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	QARKKTFSLVHEVMAVENYAKD 22
Db	25	QARKKTFSLVHEVMAVENYAKD 46

### RESULT 2

Q8WXU6	PRELIMINARY;	PRT;	138 AA.
ID Q8WXU6			
AC Q8WXU6			
DT 01-MAR-2002	(TREMBLrel. 20, Created)		
DT 01-MAR-2002	(TREMBLrel. 20, Last sequence update)		
DT 01-JUN-2003	(TREMBLrel. 24, Last annotation update)		

Db 28 AKKSNFLEPLAAHSYLAVEFYA 49

Search completed: March 18, 2004, 14:16:09  
Job time : 3.61905 secs

1000

Qy 2 ARKKTf--LSVHEVMAVENYA 20

2 ARKKTf---LSVHE'VMAVENYA 20

RR4\_RICNA STANDARD; PRT; 202 AA.  
 ID Q9M4C2;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chloroplast 30S ribosomal protein S4.  
 GN RPS4.  
 OS Ricciocarpos natans (Liverwort).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;  
 OC Marchantiopsida; Marchantiidae; Ricciales; Ricciaceae; Ricciocarpos.  
 OX NCBI\_TaxID=53035;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Gametophyte;  
 RA Capsius I.; Blocher R.;  
 RT "A molecular approach to bryophyte systematics";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds  
 CC directly to 16S rRNA where it nucleates assembly of the body of  
 CC the 30S subunit (By similarity).  
 CC -!- FUNCTION: With S5 and S12 plays an important role in translational  
 CC accuracy (By similarity).  
 CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.  
 CC The interaction surface between S4 and S5 is involved in control  
 CC of translational fidelity (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Chloroplast.  
 CC -!- SIMILARITY: Contains 1 S4 RNA-binding domain.  
 CC -!- SIMILARITY: Belongs to the S4P family of ribosomal proteins.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ251062; CAB92155.1; -  
 CC HSSP; P81288; 1C05. -  
 CC HAMAP; MF\_01306; -; 1.  
 CC InterPro; IPR001912; Ribosomal\_S4.  
 CC InterPro; IPR005709; Ribosomal\_S4\_b/o.  
 CC InterPro; IPR002942; S4.  
 CC Pfam; PF00163; Ribosomal\_S4; 1.  
 CC Pfam; PF01479; S4; 1.  
 CC SMART; SM00363; S4; 1.  
 CC TIGRFAMs; TIGR01017; rpsd\_bact; 1.  
 CC PROSITE; PS00632; RIBOSOMAL\_S4; 1.  
 CC PROSITE; PS50889; S4; 1.  
 KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.  
 FT DOMAIN 90 154 S4 RNA-BINDING.  
 SQ SEQUENCE 202 AA; 23565 MW; 607EBC24ED641FC1 CRC64;  
 Query Match 37.3%; Score 41; DB 1; Length 202;  
 Best Local Similarity 36.8%; Pred. No. 12;  
 Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 RKKTFLSVHVMVAVENYAK 21  
 DB 182 REMWFLKINELLVVEYYSR 200  
 RESULT 13  
 GSHR\_HAEIN STANDARD; PRT; 456 AA.  
 ID G3HR\_HAEIN  
 AC P43783;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-MAR-2004 (Rel. 43, Last annotation update)  
 DE Glutathione reductase (EC 1.8.1.7) (GR) (Grasee).  
 GN GOR OR H10161.

OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rd / KW20 / ATCC 51907;  
 RA Barcak G.J.; Heimer S.R.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rd / KW20 / ATCC 51907;  
 RX MEDLINE=9550630; PubMed=7542800;  
 RA Fleischmann R.D.; Adams M.D.; White O.; Clayton R.A.; Kirkness E.F.;  
 RA Kerlavage A.R.; Bult C.J.; Tomb J.-P.; Dougherty B.A.; Merrick J.M.;  
 RA McElroy K.; Sutton G.; Fitzhugh W.; Fields C.A.; Gocayne J.D.;  
 RA Scott J.D.; Shirley R.; Liu L.-I.; Glodek A.; Kelley J.M.;  
 RA Weidman J.P.; Phillips C.A.; Spriggs T.; Hedblom E.; Cotton M.D.;  
 RA Uterback T.R.; Hanna M.C.; Nguyen D.T.; Saudek D.M.; Brandon R.C.;  
 RA Fine L.D.; Fritchman J.L.; Fuhrmann J.L.; Geoghegan N.S.M.;  
 RA Gnehm C.L.; McDonald L.A.; Small K.V.; Fraser C.M.; Smith H.O.;  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT Rd".  
 RL Science 269:496-512(1995).  
 CC -!- FUNCTION: Maintain high levels of reduced glutathione in the  
 CC cytosol (By similarity).  
 CC -!- CATALYTIC ACTIVITY: 2 glutathione + NADP(+) = glutathione  
 CC disulfide + NADPH.  
 CC -!- COFACTOR: Binds 1 FAD per subunit (By similarity).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- MISCELLANEOUS: The active site is a redox-active disulfide bond.  
 CC -!- SIMILARITY: Belongs to class-I pyridine nucleotide-disulfide  
 CC oxidoreductase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U20229; AAC62137.1; -  
 CC EMBL; U32702; AAC21833.1; -  
 CC PIR; A64052; A64052.  
 CC HSSP; P00350; 1ALG.  
 CC TIGR; H10161; -  
 CC InterPro; IPR001327; FAD\_pyr\_redox.  
 CC InterPro; IPR006322; Glut\_redox\_1.  
 CC InterPro; IPR000815; Hg\_reductase.  
 CC InterPro; IPR001100; Pyr\_redox.  
 CC InterPro; IPR004099; Pyr\_redox\_dim.  
 CC InterPro; IPR000103; Pyridine\_redox\_2.  
 CC Pfam; PF00070; pyr\_redox; 1.  
 CC Pfam; PF02852; pyr\_redox\_dim; 1.  
 CC PRINTS; PR00368; FADPDR.  
 CC PRINTS; PR00945; HGRDTASE.  
 CC PRINTS; PR00411; PNDRTASEI.  
 CC PRINTS; PR00469; PNDRTASEII.  
 CC ProDom; PD000139; FAD\_pyr\_redox; 1.  
 CC TIGRFAMs; TIGR01421; Gluta\_reduc; 1.  
 CC PROSITE; PS00076; PYRIDINE\_REDOX\_1; 1.  
 KW Redox-active center; Oxidoreductase; Flavoprotein; FAD; NADP;  
 KW Complete proteome.  
 FT NP\_BIND 34 42 FAD (ADP PART) (By SIMILARITY).  
 FT DISULFID 42 47 REDOX-ACTIVE (By SIMILARITY).  
 FT ACT\_SITE 445 445 BY SIMILARITY.  
 SQ SEQUENCE 456 AA; 49329 MW; 846980C215FE949A CRC64;  
 Query Match 37.3%; Score 41; DB 1; Length 456;  
 Best Local Similarity 40.0%; Pred. No. 30;

```

DR InterPro: IPR008994; Nucleic acid OB.
DR InterPro: IPR004364; tRNA-synt_2.
DR InterPro: IPR002312; tRNA-synt_2.
DR InterPro: IPR004365; tRNA-anti.
DR InterPro: IPR006195; tRNA ligase II.
DR Pfam: PF00152; tRNA-synt_2; 1.
DR Pfam: PF01336; tRNA-anti; 1.
DR PRINTS: PR01042; tRNA synthase.
DR TIGRFAMs: TIGR00457; amsS; 1.
DR PROSITE: PS00862; AA tRNA LIGASE II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 466 AA; 54136 MW; 52946E567AD0E826 CRC64;

Query Match 38.2%; Score 42; DB 1; Length 466;
Best Local Similarity 47.1%; Pred. No. 21;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 KKTFLSVHEVMAVENVA 20
DB 204 KRTFLVSGQLHIESYA 220

RESULT 10
LON1 THEAC STANDARD; PRT; 657 AA.
AC Q9HJ89;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative protease La homolog type 1 (EC 3.4.21.-).
GN TA1081.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513(2000).
CC -1- SIMILARITY: Belongs to peptidase family S16.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AL445066; CAC12209.1; -.
CC MEROPS; S16.005; -.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR004663; Pept S16 Archaea.
DR InterPro: IPR008269; Pept S16 C.
DR InterPro: IPR001984; Peptidase_S16.
DR Pfam: PF05362; Lon_C; 1.
DR PRINTS; PR00830; ENDOLAPTASE.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00764; Lon rel; 1.
KW Hypothetical protein; Hydrolase; Serine protease; ATP-binding;
KW Complete proteome.
FT NP_BIND 57 64 ATP (POTENTIAL).
FT ACT_SITE 525 525 BY SIMILARITY.
SQ SEQUENCE 657 AA; 71600 MW; FD465CD99F5B38F7 CRC64;

Query Match 38.2%; Score 42; DB 1; Length 657;
Best Local Similarity 40.0%; Pred. No. 30;

```

```

Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 ARKKTFLSVHEVMAVENYAK 21
DB 390 SQKKTVTVTADVTAAKNLAK 409

RESULT 11
RR4 HOOLU STANDARD; PRT; 202 AA.
AC P59139;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4.
GN RPS4.
OS Hookeria lucens (Moss).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Bryidae; Hookeriales; Hookeriaceae; Hookeria.
OX NCBI_TaxID=65539;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gametophyte;
RA Blocher R., Capesius I.;
RT "The systematic position of the Hypopterygiaceae (Bryopsida) inferred
RT from rps4 gene sequences.";
RL Cryptogam. Bryol. 23:191-207(2002).
CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -1- SIMILARITY: Belongs to the S4P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AJ269689; CAC80629.1; -.
CC HAMAP; MF_01306; 1.
DR InterPro: IPR001912; Ribosomal_S4.
DR InterPro: IPR005709; Ribosomal_S4_b/o.
DR InterPro: IPR002942; S4.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRFAMs; TIGR01017; rpsD_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
DR PROSITE; PS00889; S4; 1.
KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
FT DOMAIN 90 153 S4 RNA-BINDING.
SQ SEQUENCE 202 AA; 23506 MW; C77BDB4670E62D5A CRC64;

Query Match 37.3%; Score 41; DB 1; Length 202;
Best Local Similarity 36.8%; Pred. No. 12;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAK 21
DB 182 RESTGLKINELLYVEYYSR 200

RESULT 12

```



DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Complement component C6 precursor.  
GN C6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-31 AND 633-640.  
RX MEDLINE=90036879; PubMed=2808363;  
RA Haefliger J.-A., Tschopp J., Vial N., Jenne D.E.;  
RT "Complete primary structure and functional characterization of the  
RT sixth component of the human complement system. Identification of the  
RT C5b-binding domain in complement C6.";  
RL J. Biol. Chem. 264:18041-18051(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89380223; PubMed=2789218;  
RA Discipio R.G., Hugli T.E.;  
RT "The molecular architecture of human complement component C6.";  
RL J. Biol. Chem. 264:16197-16206(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93291175; PubMed=8512929;  
RA Hobart M.J., Fernie B., Discipio R.G.;  
RT "Structure of the human C6 gene.";  
RL Biochemistry 32:6198-6205(1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Sapichenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
RP SEQUENCE OF 1-491 FROM N.A.  
RX MEDLINE=89202413; PubMed=2468158;  
RA Chakravarti D.N., Chakravarti B., Parra C.A., Mueller-Eberhard H.J.;  
RT "Structural homology of complement protein C6 with other  
RT channel-forming proteins of complement.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:2759-2803(1989).  
RN [6]  
RP VARIANT ALLOTYPES C6 A AND C6 B.  
RX MEDLINE=93326158; PubMed=8101442;  
RA Dewald G., Nothen M.M., Cichon S.;  
RT "Polymorphism of human complement component C6: an amino acid  
RT substitution (Glu/Ala) within the second thioester repeat  
RT differentiates between the two common allotypes C6 A and C6 B.";  
RL Biochem. Biophys. Res. Commun. 194:458-464(1993).  
RN [7]  
RP CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=20020247; PubMed=10551839;  
RA Hofsteenge J., Blommestein M., Hess D., Furmanek A., Miroshnichenko O.;

"The four terminal components of the complement system are C-mannosylated on multiple tryptophan residues.";  
J. Biol. Chem. 274:32786-32794(1999).  
-!- FUNCTION: Involved in the formation of the lytic c5b-9m complex.  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- PTM: ALL CYSTEINE RESIDUES ARE ASSUMED TO BE CROSS-LINKED TO ONE ANOTHER. INDIVIDUAL MODULES CONTAINING AN EVEN NUMBER OF CONSERVED CYSTEINE RESIDUES ARE SUPPOSED TO HAVE DISULFIDE LINKAGES ONLY WITHIN THE SAME MODULE.  
-!- POLYMORPHISM: The sequence shown is that of allotype C6 A.  
-!- SIMILARITY: Belongs to the complement C6/C7/C8/C9 family.  
-!- SIMILARITY: Contains 2 Sushi (SCR) domains.  
-!- SIMILARITY: Contains 3 TSP type-1 domains.  
-!- SIMILARITY: Contains 1 EGF-like domain.  
-!- SIMILARITY: Contains 1 LDL-receptor class A domain.  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; J05064; AA51860.1; -  
EMBL; J05024; AA59668.1; -  
EMBL; X72177; CAAS0994.1; -  
EMBL; BC035723; AAH35723.1; -  
EMBL; J04506; AAB59433.1; -  
PIR; A34372; A34372.  
HSSP; P01130; LAJU.  
Genew; HGNC:1339; C6.  
MIM; 217050; -  
GO; GO:0003811; F:complement activity; TAS.  
GO; GO:0006955; P:immune response; TAS.  
GO; GO:0009618; P:response to pathogenic bacteria; TAS.  
InterPro; IPR006209; EGF like.  
InterPro; IPR003884; PacI\_Mac.  
InterPro; IPR023350; kazal.  
InterPro; IPR002172; LDL\_receptor\_A.  
InterPro; IPR001862; MAC\_perforin.  
InterPro; IPR000436; Sushi\_SCR\_CCP.  
InterPro; IPR000884; TSPI.  
InterPro; IPR008085; TSP\_1.  
Pfam; PF00050; kazal; 1.  
Pfam; PF00057; ldl\_recept\_a; 1.  
Pfam; PF01823; MACPF; 1.  
Pfam; PF00084; sushi; 2.  
Pfam; PF00090; tsp; 1; 3.  
PRINTS; PR00764; COMPLEMENTC9.  
PRINTS; PR01705; TSPREPEAT.  
SMART; SM00032; CCP; 2.  
SMART; SM00057; FIMAC; 2.  
SMART; SM00280; KAZAL; 1.  
SMART; SM00192; LDLA; 1.  
SMART; SM00457; MACPF; 1.  
SMART; SM00209; TSPI; 3.  
PROSITE; PS00022; EGF\_1; 1.  
PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
PROSITE; PS01209; LDLRA\_1; 1.  
PROSITE; PS00668; LDLRA\_2; 1.  
PROSITE; PS00279; MAC\_PERFORIN; 1.  
PROSITE; PS0092; TSPI; 3.  
KW Complement pathway; Glycoprotein; Plasma; Membrane attack complex;  
KW Cytolysis; Sushi; Repeat; Signal; Polymorphism; EGF-like domain;  
KW Transmembrane.  
FT SIGNAL 1 21 COMPLEMENT COMPONENT C6.  
FT CHAIN 22 934 TSP TYPE-1 1.  
FT DOMAIN 81 134 TSP TYPE-1 2.  
FT DOMAIN 138 175 LDL-RECEPTOR CLASS A.  
FT TRANSMEM 331 349 POTENTIAL.  
FT TRANSMEM 354 373 POTENTIAL.

```

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: Orotidine 5'-phosphate =UMP + CO(2).
CC -I- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SIMILARITY: Belongs to the OMP decarboxylase family. Subfamily 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U57689; AAC02248.1; -.
DR HSSP; P25971; 1DPT.
DR HAMAP; MF 01200; -.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase; 1.
DR PROSITE; PS00156; OMPDECASE; 1.
KW Lyase; Decarboxylase; Pyrimidine biosynthesis.
FT NON TER 1
FT ACT_SITE 68 68 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 235 AA; 25093 MW; B5DB591AE0084378 CRC64;
Query Match 40.5%; Score 44.5; DB 1; Length 235;
Best Local Similarity 50.0%; Pred. No. 3.7;
Matches 14; Conservative 1; Mismatches 6; Indels 7; Gaps 2;
QY 1 QARKTKFLSV-----HEV-MAVENYAK 21
DQ ||||| : | | | | |
DB 58 QARKTKVFFDMKGLDIDHTVARAVENYAK 85
RESULT 5
YLP3_PSEPU STANDARD; PRT; 394 AA.
AC P31049;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 44.7 kDa protein in lpd-3 5' region (ORF3).
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G2;
RA Lorenz D., Sokatch J.R.;
RA Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X55704; CAA39234.1; -.
DR InterPro; IPR003333; CMAS.
DR InterPro; IPR001601; Methyitransf.
DR Pfam; PF02353; CMAS; 1.
KW Hypothetical protein.
SQ SEQUENCE 394 AA; 44656 MW; 8CD61FE9495302AF CRC64;
Query Match 39.1%; Score 43; DB 1; Length 394;
Best Local Similarity 36.8%; Pred. No. 12;
Matches 7; Conservative 9; Mismatches 3; Indels 0; Gaps 0;
QY 2 ARKTKFLSVHEVMAVENYA 20

```

```

Db 361 AFEKGWINLHQILAVKPYA 379
RESULT 6
LBD2_ARATH STANDARD; PRT; 205 AA.
AC Q9LNB9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative LOB domain protein 2.
GN LBD2 OR ATG06280 OR F9P14.14.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RC MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin B.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Taiton L.J., Tambunga G., Toriumi M.J., Town C.D., Walker M.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
CC -I- SIMILARITY: Contains 1 LOB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AC052590; AAF80225.1; -.
DR PIR; F86198; F86198.
DR InterPro; IPR004883; DUF260.
DR Pfam; PF03195; DUF260; 1.
DR PROSITE; PS0891; LOB; 1.
KW Hypothetical protein.
FT DOMAIN 22 122 LOB.
SQ SEQUENCE 205 AA; 23285 MW; 404FD0C4FC32A342 CRC64;
Query Match 38.6%; Score 42.5; DB 1; Length 205;
Best Local Similarity 52.4%; Pred. No. 7;
Matches 11; Conservative 2; Mismatches 7; Indels 1; Gaps 1;
QY 2 ARKTKFLSVHEVMAVENYAK 21
DQ ||||| : | | | | |
DB 46 ARKTKFQAVHKVFGVSNVQK 66
RESULT 7
CO6 HUMAN STANDARD; PRT; 934 AA.
ID CO6 HUMAN
AC P13671;

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR ENBL; D42125; BAA07705.1; -.  
 DR ENBL; AE003784; AAM68337.1; -.  
 DR ENBL; AY058652; AAL13881.1; -.  
 DR ENBL; S55977; AAB19907.1; -.  
 DR ENBL; AJ002911; CAA05746.1; -.  
 DR HSSP; P00523; 1SRL.  
 DR FlyBase; FBgn0004603; Src42A.  
 DR GO; GO:0005912; C-adherens junction; IDA.  
 DR GO; GO:0004713; P-protein-tyrosine kinase activity; NAS.  
 DR GO; GO:0007154; P-cell communication; IDA.  
 DR GO; GO:0007456; P-eye morphogenesis (sensu Drosophila); IMP.  
 DR GO; GO:0007169; P-transmembrane receptor protein tyrosine kin. .; IGI.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_kinase\_AS.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; pkinaase; 1.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 KW Transferase; Tyrosine-protein kinase; Developmental protein;  
 KW ATP-binding; SH2 domain; SH3 domain.  
 FT DOMAIN 63 124  
 FT D0MAIN 130 222 SH2.  
 FT D0MAIN 248 504 PROTEIN\_KINASE.  
 FT NP\_BIND 254 262 ATP (BY SIMILARITY).  
 FT BINDING 276 276 ATP (BY SIMILARITY).  
 FT ACT\_SITE 370 370 BY SIMILARITY.  
 FT CONFLICT 65 65 A -> V (IN REF. 1).  
 FT CONFLICT 376 376 V -> I (IN REF. 4 AND 5).  
 FT CONFLICT 381 386 GNIVKI -> SNVVKM (IN REF. 4).  
 SQ SEQUENCE 517 AA; 59069 MW; 1EF196B4D7AE61E9 CRC64;  
 Query Match 40.9%; Score 45; DB 1; Length 517;  
 Best Local Similarity 47.6%; Pred. No. 7.1;  
 Matches 10; Conservative 6; Mismatches 3; Indels 2; Gaps 1;  
 QY 2 ARKKTPLSVHEWAVENYAKD 22  
 DB 192 ARRTFTRLQEL--VEHYSKD 210  
 RESULT 4  
 PYRF\_BARBA  
 ID PYRF\_BARBA STANDARD; PRT; 235 AA.  
 AC Q44843;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)  
 DE (OMPDCase) (OMPDcase) (Fragment).  
 PYRF.  
 GN GN  
 OS Bartonella bacilliformis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bartonellaceae; Bartonella.  
 OX NCBI\_TaxID=774;  
 RN [1]  
 RA SEQUENCE FROM N.A.  
 RP Upeslaciis E., Ihler G.M.;

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AL096677; CAC13170.1; -  
EMBL; AL096677; CAC17423.1; -  
HSSP; P01038; 1A90.  
Genew; HGNC:15959; CSTR11.  
InterPro; IPR000010; Cystatin.  
Pfam; PF00031; cystatin; 1.  
SMART; SM00043; CY; 1.  
PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
KW Thiol protease inhibitor; Signal; Alternative splicing.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 137 CYPSTATIN 11.  
FT SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).  
FT DISULFID 93 101 BY SIMILARITY.  
FT DISULFID 114 134 BY SIMILARITY.  
FT CARBOHYD 131 131 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT VARSPLIC 76 110 Missing (in isoform 2).  
FTid=VSP\_001260.  
SQ SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 110; DB 1; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.1e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QARKTFLSVHEVMAVENYAKD 22  
DB 24 QARKTFLSVHEVMAVENYAKD 45

## RESULT 2

CS11 MOUSE STANDARD; PRT; 139 AA.  
AC Q9D285;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cystatin 11 precursor.  
GN CST11.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Epithelium;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Guncich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Hayashaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuke S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -1- SIMILARITY: Belongs to the cystatin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AK020300; BAB32061.1; -  
HSSP; P01034; 1G96.  
MGI; MGI:1925490; Cst11.  
InterPro; IPR000010; Cystatin.  
Pfam; PF00031; cystatin; 1.  
SMART; SM00043; CY; 1.  
PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
KW Thiol protease inhibitor; Signal.  
FT SIGNAL 1 28 POTENTIAL.  
FT CHAIN 29 139 CYPSTATIN 11.  
FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).  
FT DISULFID 94 102 BY SIMILARITY.  
FT DISULFID 115 135 BY SIMILARITY.  
FT CARBOHYD 134 134 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 139 AA; 16217 MW; F228D9815PA32640 CRC64;

Query Match 48.2%; Score 53; DB 1; Length 139;  
Best Local Similarity 40.9%; Pred. No. 0.071;  
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 QARKTFLSVHEVMAVENYAKD 22  
DB 25 QVKKRTFIRIEVSALSSVKE 46

## RESULT 3

SR42 DROME STANDARD; PRT; 517 AA.  
ID SR42 DROME  
AC Q9VJ3; O18369; Q26297; Q94879;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DE 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tyrosine-protein kinase Src42A (EC 2.7.1.112) (Darc41).  
GN SRC42A OR SRC41 OR TK5 OR CG7873  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.  
RC STRAIN=Canton-S; TISSUE=Pupa;  
RX MEDLINE=96268448; PubMed=8692295;  
RA Takahashi F., Endo S., Kojima T., Saigo K.;  
RT "Regulation of cell-cell contacts in developing Drosophila eyes by Darc41, a new, close relative of vertebrate c-src."  
RL Genes Dev. 10:1645-1656(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champ M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 2.61905 Seconds  
(without alignments)  
437.389 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKKTFLSVHEVWVAVENYAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	137	CS11 HUMAN	Q9h112 homo sapien
2	53	48.2	139	CS11 MOUSE	Q9d269 mus musculu
3	45	40.9	517	SR42 DROME	Q9v913 drosophila
4	44.5	40.5	235	YPRF BARBA	Q44843 bartonella
5	43	39.1	394	YLP3 PSEPU	P31049 pseudomonas
6	42.5	38.6	205	1 LBD2 ARATH	Q9lnb9 arabidopsis
7	42.5	38.6	934	CO6 HUMAN	P13671 homo sapien
8	42.5	38.6	1056	1 K125 ARATH	P82266 arabidopsis
9	42	38.2	466	1 SYN WIGBR	Q8d2u1 wigglewort
10	42	38.2	657	1 LONI THEAC	Q9hj89 thermoplasm
11	41	37.3	202	1 RR4 HOOLU	P59139 hookeria lu
12	41	37.3	202	1 RR4 RICNA	Q9m4c2 ricciocarpo
13	41	37.3	456	1 GSHR HABIN	P43783 haemophilus
14	40.5	36.8	264	1 SPED PSEAE	Q91577 pseudomonas
15	40.5	36.8	508	1 GSBP CHICK	P12244 gallus gall
16	40	36.4	237	1 PRI2 SOYBN	Q94ic4 glycine max
17	40	36.4	268	1 LBD2 ARATH	Q9lrv1 arabidopsis
18	40	36.4	292	1 YZ51 MYCTU	P71850 mycobacteri
19	40	36.4	294	1 SUCD METJA	Q58643 methanococ
20	40	36.4	437	1 SUCB NEOPR	P53977 neocallima
21	40	36.4	496	1 Y048 UREPA	Q9pr97 ureaplasma
22	40	36.4	977	1 UB24 HUMAN	Q9upus homo sapien
23	40	36.4	6629	1 RIAB IBVBC	P27920 a replicase
24	40	36.4	6629	1 RIAB IBVBC	Q91qt2 a replicase
25	39.5	35.9	264	1 SPED PSESM	Q889x9 pseudomonas
26	39	35.5	202	1 RR4 HYPLA	P59145 hypoterygi
27	39	35.5	249	1 CORJ METJA	Q58223 methanococ
28	39	35.5	310	1 YGRJ YEAST	P53108 saccharomyc
29	39	35.5	382	1 DHGY CUCSA	P13443 caccumis sat
30	39	35.5	398	1 DXR PHOLI	Q7n8p3 photorhabdu
31	39	35.5	431	1 AROA AQUAE	O67494 aquifex aeo
32	39	35.5	466	1 SYN XYLFA	Q9paf5 xylella fas
33	39	35.5	466	1 SYN_XYLFT	Q87a82 xylella fas

34	39	35.5	469	1	1A13 LYCES	Q42881 lycopersico
35	39	35.5	690	1	CAN9 HUMAN	O14815 homo sapien
36	39	35.5	888	1	YJH0 YEAST	P40361 saccharomyc
37	38.5	35.0	610	1	RYK1 DROME	Q27324 drosophila
38	38.5	35.0	1523	1	DPOL THEFM	P74918 thermococcu
39	38	34.5	133	1	PABP SCHMA	P29498 schistosoma
40	38	34.5	165	1	LB21 ARATH	O59r18 arabidopsis
41	38	34.5	185	1	RR4 WOOPA	P59151 woodwardia
42	38	34.5	202	1	RR4 MARPO	P06358 marchantia
43	38	34.5	202	1	RR4 PLARP	Q9m4c3 plagiocnasm
44	38	34.5	205	1	RR4 MESVI	Q9mun0 mesostigma
45	38	34.5	270	1	CNX3 ARATH	Q239056 arabidopsis

## ALIGNMENTS

### RESULT 1

ID	CS11_HUMAN	STANDARD;	PRT;	137 AA.
AC	Q9H112; Q9H113;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cystatin 11 precursor.			
GN	CST11 OR CST8L			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RK	MEDLINE=21638749; PubMed=11780052;			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Corby N.R.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,			
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,			
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,			
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,			
RA	Key M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,			
RA	Lehvaeslaih M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,			
RA	Marsh V.L., Martin S.L., McConachie L.J., McMay K., McMurray A.A.,			
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,			
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,			
RA	Phillimore B.J.C.T., Prachalam S.R., Plumb R.W., Ramsay H.,			
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,			
RA	Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,			
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,			
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,			
RA	Whithead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,			
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,			
RA	Rogers J.;			
RT	"The DNA sequence and comparative analysis of human chromosome 20.";			
RL	Nature 414:865-871(2001).			
CC	-1- SUBCELLULAR LOCATION: Secreted (Potential).			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=1;			
CC	isoId=Q9H112-1; Sequence=Displayed;			
CC	Name=2;			
CC	isoId=Q9H112-2; Sequence=VSP_001260;			
CC	Note=No experimental confirmation available;			
CC	-1- SIMILARITY: Belongs to the cystatin family.			

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its



A;Accession: H84777  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1056 <STO>  
 A;Cross-references: GB:AE002093; NID:g4510356; PIDN:AAD21445.1; GSPDB:GN00139  
 C;Genetics:  
 A;Gene: At2g36200  
 A;Map position: 2  
 C;Superfamily: kinesin-related protein Eg5; kinesin motor domain homology

Query Match 38.6%; Score 42.5; DB 2; Length 1056;  
 Best Local Similarity 61.1%; Pred. No. 1.3e+02;  
 Matches 11; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 5 KTFLSVHEVMVAVENYAKD 22  
 ||||| | | | | | | | | | |  
 Db 770 KTFLEDEH-VSAVNNLTND 786

RESULT 15  
 D90026  
 hypothetical protein [imported] - Staphylococcus aureus (strain N315)  
 C;Species: Staphylococcus aureus  
 C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C;Accession: D90026  
 R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A;Reference number: A89758; MUID:21311952; PMID:11418146  
 A;Accession: D90026  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-256 <KUR>  
 A;Cross-references: GB:BA000018; PID:g13702081; PIDN:BA043373.1; GSPDB:GN00149  
 A;Experimental source: strain N315  
 C;Genetics:  
 A;Gene: SA2076

Query Match 38.2%; Score 42; DB 2; Length 256;  
 Best Local Similarity 44.4%; Pred. No. 35;  
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 KTFLSVHEVMVAVENYAKD 22  
 | : | | | : | |  
 Db 142 KDYADAHEEMVREHYQKD 159

Search completed: March 18, 2004, 14:23:13  
 Job time : 5.23413 secs



Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 4 KKTFLSVHEVMAVENYAKD 22  
|||:::||:|::|

Db 328 KKVFKSLHNPTIILYSKE 346  
|||:::||:|::|

RESULT 7

A81037 tyrosyl-tRNA synthetase NMB1835 [imported] - Neisseria meningitidis (strain MC58 serogroup B)

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001

C:Accession: A81037

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickley, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; PMID:20175755; PMID:10710307

A:Accession: A81037

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-431 <TET>

A:Cross-references: GB:AE002534; GB:AE002098; NID:G7227095; PIDN:AAF42170.1; PID:G7227095

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB1835

C:Superfamily: tyrosine-tRNA ligase

Query Match 39.1%; Score 43; DB 2; Length 431;  
Best Local Similarity 58.8%; Pred. No. 41;  
Matches 10; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

Qy 6 TFLSVHEVMAVENYAKD 22  
||||:|::|::|

Db 272 TFLSIEIDAIE--AKD 286  
||||:|::|::|

RESULT 8

H81981 probable tyrosine-tRNA ligase (EC 6.1.1.1) NMA0620 [imported] - Neisseria meningitidis

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 03-Jun-2002

C:Accession: H81981

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; PMID:20222556; PMID:10761919

A:Accession: H81981

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-431 <PAR>

A:Cross-references: GB:ALI62753; GB:ALI157959; NID:G7379120; PIDN:CAB83910.1; PID:G737935

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: tyrS; NMA0620

C:Superfamily: tyrosine-tRNA ligase

C:Keywords: ligase

Query Match 39.1%; Score 43; DB 2; Length 431;  
Best Local Similarity 58.8%; Pred. No. 41;  
Matches 10; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

Qy 6 TFLSVHEVMAVENYAKD 22  
||||:|::|::|

Db 272 TFLSIEIDAIE--AKD 286  
||||:|::|::|

RESULT 9

T15892 hypothetical protein D2096.4 - Caenorhabditis elegans



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 4.23413 Seconds  
(without alignments)  
499.799 Million cell updates/sec

Title: US-09-941-314-5  
Perfect score: 110  
Sequence: 1 QARKKTFLSVHEWVAENVAKD 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	48.2	184	2	D72061
2	53	48.2	184	2	B86562
3	45	40.9	706	2	S61717
4	43	39.1	102	2	T07890
5	43	39.1	259	2	F64575
6	43	39.1	366	2	T21166
7	43	39.1	431	2	A81037
8	43	39.1	431	2	H81981
9	43	39.1	467	2	T15892
10	43	39.1	585	2	G96995
11	43	39.1	676	2	G84663
12	42.5	38.6	205	2	F86398
13	42.5	38.6	934	1	A43372
14	42.5	38.6	1056	2	H84777
15	42	38.2	256	2	D90026
16	42	38.2	271	2	E84138
17	41	37.3	216	2	E87274
18	41	37.3	280	2	C86889
19	41	37.3	297	2	AH2358
20	41	37.3	305	2	D83488
21	41	37.3	326	2	F71808
22	41	37.3	423	2	F64690
23	41	37.3	429	2	G69368
24	41	37.3	456	2	A64052
25	41	37.3	769	2	D86678
26	41	37.3	1211	2	T23210
27	41	37.3	1422	2	T18404
28	40.5	36.8	264	2	D83564
29	40.5	36.8	508	1	A30007

30	40	36.4	292	2	H70677
31	40	36.4	294	2	E64455
32	40	36.4	387	2	A90134
33	40	36.4	414	2	T19912
34	40	36.4	426	2	D72252
35	40	36.4	496	2	E82940
36	40	36.4	662	2	D83367
37	40	36.4	934	2	T47546
38	40	36.4	949	2	F84779
39	40	36.4	1088	2	T05464
40	40	36.4	2652	1	VFIHB2
41	39	35.5	169	2	T32536
42	39	35.5	212	2	G90353
43	39	35.5	249	2	E64401
44	39	35.5	298	2	T06980
45	39	35.5	310	2	S60425

## ALIGNMENTS

### RESULT 1

D72061  
hypothetical protein CP0175 [imported] - Chlamydothila pneumoniae (strains CWL029 and AR-  
C;Species: Chlamydothila pneumoniae, Chlamydia pneumoniae  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
R;Accession: D72061; A81606  
R;Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999  
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A;Reference number: A72000; MUID:99206606; PMID:10192388  
A;Accession: D72061  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-184 <ARN>  
A;Cross-references: GB:AE001642; GB:AE001363; NID:g4376865; PIDN:AAD18714.1; PID:g4376865  
A;Experimental source: strain CWL029  
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, F.;  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A;Reference number: A81500; MUID:20150255; PMID:10684935  
A;Accession: A81606  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-184 <REA>  
A;Cross-references: GB:AE002178; GB:AE002161; NID:g7189099; PIDN:AAF38049.1; PID:g7189101  
A;Experimental source: strain AR39, HL cells  
C;Genetics:  
A;Gene: CP00574; CP0175

Query Match 48.2%; Score 53; DB 2; Length 184;  
Best Local Similarity 55.0%; Pred. No. 0.39;  
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEWVAENVAKD 22  
|||||||:|:|:|  
DB 26 RKKTFLSHRVLARPSFPVD 45

### RESULT 2

B86562  
hypothetical protein CPJ0574 [imported] - Chlamydothila pneumoniae (strain J138)  
C;Species: Chlamydothila pneumoniae, Chlamydia pneumoniae  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise  
Nucleic Acids Res. 28, 2311-2314, 2000  
A;Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.  
A;Reference number: A86491; MUID:20330349; PMID:10871362  
A;Accession: B86562  
A;Status: preliminary  
A;Molecule type: DNA



Db 105 KALLSVHDTVAKNY 119

RESULT 13

US-09-783-320-22  
 ; Sequence 22, Application US/09783320  
 ; Patent No. US20020038011A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Walke, D. Wade  
 ; APPLICANT: Hu, Yi  
 ; APPLICANT: Nepomnichy, Boris  
 ; APPLICANT: Turner, C. Alexander Jr  
 ; APPLICANT: Zambrowicz, Brian  
 ; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding  
 ; FILE REFERENCE: LEX-0137-USA  
 ; CURRENT APPLICATION NUMBER: US/09/783,320  
 ; CURRENT FILING DATE: 2001-02-15  
 ; PRIOR APPLICATION NUMBER: US 60/183,582  
 ; PRIOR FILING DATE: 2000-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/184,014  
 ; PRIOR FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 22  
 ; LENGTH: 473  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-09-783-320-22

Query Match 40.0%; Score 44; DB 9; Length 473;  
 Best Local Similarity 53.3%; Pred. No. 84;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 KTFLSVHEVMAVENY 19  
 | ||||: :||  
 Db 105 KALLSVHDTVAKNY 119

RESULT 14

US-09-783-320-30  
 ; Sequence 30, Application US/09783320  
 ; Patent No. US20020038011A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Walke, D. Wade  
 ; APPLICANT: Hu, Yi  
 ; APPLICANT: Nepomnichy, Boris  
 ; APPLICANT: Turner, C. Alexander Jr  
 ; APPLICANT: Zambrowicz, Brian  
 ; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding  
 ; FILE REFERENCE: LEX-0137-USA  
 ; CURRENT APPLICATION NUMBER: US/09/783,320  
 ; CURRENT FILING DATE: 2001-02-15  
 ; PRIOR APPLICATION NUMBER: US 60/183,582  
 ; PRIOR FILING DATE: 2000-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/184,014  
 ; PRIOR FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 30  
 ; LENGTH: 520  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-09-783-320-30

Query Match 40.0%; Score 44; DB 9; Length 520;  
 Best Local Similarity 53.3%; Pred. No. 94;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 KTFLSVHEVMAVENY 19  
 | ||||: :||  
 Db 105 KALLSVHDTVAKNY 119

RESULT 15

US-10-094-749-3093  
 ; Sequence 3093, Application US/10094749  
 ; Publication No. US20030219741A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ISOGAI, TAKAO  
 ; APPLICANT: SUGIYAMA, TOMOYASU  
 ; APPLICANT: OTSUKI, TETSUJI  
 ; APPLICANT: WAKAMATSU, AI  
 ; APPLICANT: SATO, HIROYUKI  
 ; APPLICANT: ISHII, SHIZUKO  
 ; APPLICANT: YAMAMOTO, JUN-ICHI  
 ; APPLICANT: ISONO, YUUKO  
 ; APPLICANT: HIO, YURI  
 ; APPLICANT: OTSUKA, KAORU  
 ; APPLICANT: NAGAI, KEIICHI  
 ; APPLICANT: IRIE, RYOTARO  
 ; APPLICANT: TAMECHIKA, ICHIRO  
 ; APPLICANT: SEKI, NAOHIKO  
 ; APPLICANT: YOSHIKAWA, TSUTOMU  
 ; APPLICANT: OTSUKA, MOTOYUKI  
 ; APPLICANT: NAGAHARI, KENJI  
 ; APPLICANT: MASUHO, YASUHIKO  
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
 ; FILE REFERENCE: 084335/0160  
 ; CURRENT APPLICATION NUMBER: US/10/094,749  
 ; CURRENT FILING DATE: 2002-03-12  
 ; PRIOR APPLICATION NUMBER: 60/350,435  
 ; PRIOR FILING DATE: 2002-01-24  
 ; PRIOR APPLICATION NUMBER: JP 2001-328381  
 ; PRIOR FILING DATE: 2001-09-14  
 ; NUMBER OF SEQ ID NOS: 3381  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3093  
 ; LENGTH: 542  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-094-749-3093

Query Match 40.0%; Score 44; DB 15; Length 542;  
 Best Local Similarity 53.3%; Pred. No. 98;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 KTFLSVHEVMAVENY 19  
 | ||||: :||  
 Db 105 KALLSVHDTVAKNY 119

Search completed: March 18, 2004, 14:30:52  
 Job time : 13.3135 secs

SOFTWARE: SeqWin99, version 1.02  
; SEQ ID NO 375  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-10-312-273-375

Query Match 48.2%; Score 53; DB 15; Length 184;  
Best Local Similarity 55.0%; Pred. No. 0.91;  
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22  
Db 26 RKKTFLSSHRVLARPPVD 45

## RESULT 9

US-10-289-762-616  
; Sequence 616, Application US/10289762  
; Publication No. US20040006219A1  
; GENERAL INFORMATION:  
; APPLICANT: Grifais, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/10/289,762  
; CURRENT FILING DATE: 2003-03-27  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 616  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-10-289-762-616

Query Match 44.5%; Score 49; DB 15; Length 195;  
Best Local Similarity 50.0%; Pred. No. 4.5;  
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22  
Db 37 RQKTFLSSHRVLARPPVD 56

## RESULT 10

US-09-783-320-8  
; Sequence 8, Application US/09783320  
; Patent No. US20020038011A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Hu, Yi  
; APPLICANT: Nepomnichy, Boris  
; APPLICANT: Turner, C. Alexander Jr  
; APPLICANT: Zambrowicz, Brian  
; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding  
; FILE REFERENCE: LEX-0137-USA  
; CURRENT APPLICATION NUMBER: US/09/783,320  
; CURRENT FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: US 60/183,582  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: US 60/184,014  
; PRIOR FILING DATE: 2000-02-22  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 296  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-783-320-8

Query Match 40.08; Score 44; DB 9; Length 296;  
Best Local Similarity 53.3%; Pred. No. 49;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 KTFSLSVHEVMAVENY 19  
Db 105 KALLSVHDTVAQKNY 119

## RESULT 11

US-09-783-320-12  
; Sequence 12, Application US/09783320  
; Patent No. US20020038011A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Hu, Yi  
; APPLICANT: Nepomnichy, Boris  
; APPLICANT: Turner, C. Alexander Jr  
; APPLICANT: Zambrowicz, Brian  
; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding  
; FILE REFERENCE: LEX-0137-USA  
; CURRENT APPLICATION NUMBER: US/09/783,320  
; CURRENT FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: US 60/183,582  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: US 60/184,014  
; PRIOR FILING DATE: 2000-02-22  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-783-320-12

Query Match 40.0%; Score 44; DB 9; Length 318;  
Best Local Similarity 53.3%; Pred. No. 53;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 KTFSLSVHEVMAVENY 19  
Db 105 KALLSVHDTVAQKNY 119

## RESULT 12

US-09-783-320-18  
; Sequence 18, Application US/09783320  
; Patent No. US20020038011A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Hu, Yi  
; APPLICANT: Nepomnichy, Boris  
; APPLICANT: Turner, C. Alexander Jr  
; APPLICANT: Zambrowicz, Brian  
; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding  
; FILE REFERENCE: LEX-0137-USA  
; CURRENT APPLICATION NUMBER: US/09/783,320  
; CURRENT FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: US 60/183,582  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: US 60/184,014  
; PRIOR FILING DATE: 2000-02-22  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-783-320-18

Query Match 40.0%; Score 44; DB 9; Length 375;  
Best Local Similarity 53.3%; Pred. No. 64;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 KTFSLSVHEVMAVENY 19  
Db 105 KALLSVHDTVAQKNY 119

; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 34822  
 ; LENGTH: 50  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AL109954.10  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5  
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6  
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1  
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2  
 ; OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23  
 ; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01  
 ; US-09-864-761-34822

Query Match 91.8%; Score 101; DB 9; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22  
 |||||  
 DB 1 RKKTFLSVHEVMAVENYAKD 20

RESULT 7

US-09-864-761-48936  
 ; Sequence 48936, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharon G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; FILE REFERENCE: Aeomica-x-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 48936  
 ; LENGTH: 50  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AL096677.18  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96  
 ; OTHER INFORMATION: EST\_HUMAN HIT: A1200857.1, EVALUATE 5.00e-23  
 ; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01  
 ; US-09-864-761-48936

Query Match 91.8%; Score 101; DB 9; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22  
 |||||  
 DB 1 RKKTFLSVHEVMAVENYAKD 20

RESULT 8

US-10-312-273-375  
 ; Sequence 375, Application US/10312273  
 ; Publication No. US20040005667A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHIRON SPA  
 ; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE  
 ; FILE REFERENCE: P025035WO  
 ; CURRENT APPLICATION NUMBER: US/10/312,273  
 ; CURRENT FILING DATE: 2002-12-20  
 ; PRIOR APPLICATION NUMBER: 0016363.4  
 ; PRIOR FILING DATE: 2000-07-03  
 ; PRIOR APPLICATION NUMBER: 0017047.2  
 ; PRIOR FILING DATE: 2000-07-11  
 ; PRIOR APPLICATION NUMBER: 0017983.8  
 ; PRIOR FILING DATE: 2000-07-21  
 ; PRIOR APPLICATION NUMBER: 0019368.0  
 ; PRIOR FILING DATE: 2000-08-07  
 ; PRIOR APPLICATION NUMBER: 0020440.4  
 ; PRIOR FILING DATE: 2000-08-18  
 ; PRIOR APPLICATION NUMBER: 0022583.9  
 ; PRIOR FILING DATE: 2000-09-14  
 ; PRIOR APPLICATION NUMBER: 0027549.5  
 ; PRIOR FILING DATE: 2000-11-10  
 ; PRIOR APPLICATION NUMBER: 0031706.5  
 ; PRIOR FILING DATE: 2000-12-22  
 ; NUMBER OF SEQ ID NOS: 664

; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941.314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-3

Query Match 100.0%; Score 110; DB 9; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1.8e-10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QARKKTFSLSVHEVMAVENYAKD 22  
Db 2 QARKKTFSLSVHEVMAVENYAKD 23

## RESULT 3

US-09-941-314-4  
; Sequence 4, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941.314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-4

Query Match 100.0%; Score 110; DB 9; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1.8e-10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QARKKTFSLSVHEVMAVENYAKD 22  
Db 4 QARKKTFSLSVHEVMAVENYAKD 25

## RESULT 4

US-09-941-314-2  
; Sequence 2, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941.314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 137  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-2

Query Match 100.0%; Score 110; DB 9; Length 137;  
Best Local Similarity 100.0%; Pred. No. 2.2e-10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QARKKTFSLSVHEVMAVENYAKD 22  
Db 24 QARKKTFSLSVHEVMAVENYAKD 45

## RESULT 5

US-09-941-314-6  
; Sequence 6, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941.314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-6

Query Match 91.8%; Score 101; DB 9; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RKKTFSLSVHEVMAVENYAKD 22  
Db 1 RKKTFSLSVHEVMAVENYAKD 20

## RESULT 6

US-09-864-761-34822  
; Sequence 34822, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 13.3135 Seconds  
(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKKTFSLVHEVMAVENYAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	22	9 US-09-941-314-5	Sequence 5, Appli
2	110	100.0	115	9 US-09-941-314-3	Sequence 3, Appli
3	110	100.0	117	9 US-09-941-314-4	Sequence 4, Appli
4	110	100.0	137	9 US-09-941-314-2	Sequence 2, Appli
5	101	91.8	36	9 US-09-941-314-6	Sequence 6, Appli
6	101	91.8	50	9 US-09-864-761-34822	Sequence 34822, A
7	101	91.8	50	9 US-09-864-761-48936	Sequence 48936, A
8	53	48.2	184	15 US-10-312-273-375	Sequence 375, App
9	49	44.5	195	15 US-10-289-762-616	Sequence 616, App
10	44	40.0	296	9 US-09-783-320-8	Sequence 8, Appli
11	44	40.0	318	9 US-09-783-320-12	Sequence 12, Appli
12	44	40.0	375	9 US-09-783-320-18	Sequence 18, Appli
13	44	40.0	473	9 US-09-783-320-22	Sequence 22, Appli
14	44	40.0	520	9 US-09-783-320-30	Sequence 30, Appli
15	44	40.0	542	15 US-10-094-749-3093	Sequence 3093, App

16	44	40.0	560	9 US-09-783-320-38	Sequence 38, Appli
17	44	40.0	576	9 US-09-783-320-46	Sequence 46, Appli
18	44	40.0	576	9 US-09-994-288-5	Sequence 5, Appli
19	44	40.0	576	16 US-10-311-034-16	Sequence 16, Appli
20	43	39.1	107	12 US-10-424-599-219730	Sequence 219730, A
21	43	39.1	259	9 US-09-895-913A-338	Sequence 338, App
22	43	39.1	431	12 US-10-282-122A-65756	Sequence 65756, A
23	43	39.1	467	15 US-10-347-470A-4	Sequence 4, Appli
24	43	39.1	1054	9 US-09-815-242-4995	Sequence 4995, Ap
25	43	39.1	1175	12 US-10-282-122A-57562	Sequence 57562, A
26	43	39.1	1208	9 US-09-815-242-10627	Sequence 10627, A
27	43	39.1	1217	12 US-10-282-122A-42405	Sequence 42405, A
28	42.5	38.6	281	12 US-10-282-122A-61171	Sequence 61171, A
29	42	38.2	62	12 US-10-424-599-182304	Sequence 182304, A
30	42	38.2	212	12 US-10-425-114-38286	Sequence 38286, A
31	42	38.2	273	12 US-10-424-599-227071	Sequence 227071, A
32	42	38.2	437	12 US-10-282-122A-61087	Sequence 61087, A
33	42	38.2	577	12 US-10-425-114-57760	Sequence 57760, A
34	42	38.2	597	12 US-10-424-599-207456	Sequence 207456, A
35	41	37.3	214	12 US-10-424-599-182326	Sequence 182326, A
36	41	37.3	244	12 US-10-424-599-157243	Sequence 157243, A
37	41	37.3	455	12 US-10-169-103-2	Sequence 2, Appli
38	41	37.3	456	12 US-10-282-122A-58093	Sequence 58093, A
39	41	37.3	456	14 US-10-260-877-78	Sequence 78, Appli
40	40.5	36.8	463	12 US-10-425-114-64141	Sequence 64141, A
41	40	36.4	65	12 US-10-424-599-254688	Sequence 254688, A
42	40	36.4	86	12 US-10-424-599-199963	Sequence 199963, A
43	40	36.4	102	12 US-10-424-599-177539	Sequence 177539, A
44	40	36.4	119	12 US-10-424-599-244889	Sequence 244889, A
45	40	36.4	185	12 US-10-424-599-268671	Sequence 268671, A

ALIGNMENTS

RESULT 1

US-09-941-314-5  
; Sequence 5, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-5

Query Match 100.0%; Score 110; DB 9; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2.7e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	QARKKTFSLVHEVMAVENYAKD 22
Db	1	QARKKTFSLVHEVMAVENYAKD 22

RESULT 2

US-09-941-314-3  
; Sequence 3, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein



```
; EARLIER APPLICATION NUMBER: 60/134,175
; EARLIER FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Human
US-09-422-869-28
```

```
Query Match      35.5%; Score 39; DB 3; Length 690;
Best Local Similarity 53.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 QARKKTPLSVHEV 13
      :||| |||::||
Db      443 RARSKTFINLREV 455
```

```
Search completed: March 18, 2004, 14:25:35
Job time : 6.63095 secs
```

US-09-328-352-6399

Query Match 35.5%; Score 39; DB 4; Length 291;  
Best Local Similarity 33.3%; Pred. No. 1.1e+02;  
Matches 7; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

QY 3 RKKTFLSVHEVM--AVENYAK 21  
DB 125 RREELGIHQALAGAVDSYAR 145

RESULT 12

US-09-134-000C-5712  
; Sequence 5712, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/134,000C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/055,778

; NUMBER OF SEQ ID NOS: 6812

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5712

; LENGTH: 399

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-134-000C-5712

Query Match 35.5%; Score 39; DB 4; Length 399;  
Best Local Similarity 47.4%; Pred. No. 1.5e+02;  
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 4 KKTFLSVHEVMVAVENYAKD 22  
DB 125 KSTLQTQVQEPMAKTPFAKD 143

RESULT 13

US-08-378-313-33

; Sequence 33, Application US/08378313

; Patent No. 6207881

; GENERAL INFORMATION:

; APPLICANT: THEOLOGIS, ATHANASIOS

; APPLICANT: SATO, TAKAHIDO

; TITLE OF INVENTION: CONTROL OF FRUIT RIPENING THROUGH

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON &amp; FOERSTER

; STREET: 755 Page Mill Road

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94304-1018

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/378,313

; FILING DATE:

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/862,493

; FILING DATE: 02-APR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: MURASHIGE, KATE H.

; REGISTRATION NUMBER: 29,959

; REFERENCE/DOCKET NUMBER: 29190-20002.20

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 856-5600

; TELEFAX: (415) 494-0792

; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 469 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-378-313-33

Query Match

Best Local Similarity

Matches 8; Conservative

35.5%; Score 39; DB 3; Length 469;

Pred. No. 1.8e+02;

Mismatches 4; Indels 0; Gaps 0;

QY 7 FLSVHEVMVAVENYAK 21

DB 243 FVSVMEVLEIKNYMK 257

RESULT 14

US-09-377-497-70

; Sequence 70, Application US/09377497

; Patent No. 6670119

; GENERAL INFORMATION:

; APPLICANT: YOSHIKAWA, YOSHIE

; APPLICANT: MUKAI, HIROYUKI

; APPLICANT: ASADA, KIYOZO

; APPLICANT: HINO, FUMITSUGU

; APPLICANT: KATO, IKUNOSHIN

; TITLE OF INVENTION: CANCER-ASSOCIATED GENES

; FILE REFERENCE: 1422-388P

; CURRENT APPLICATION NUMBER: US/09/377,497

; CURRENT FILING DATE: 1999-08-20

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 70

; LENGTH: 664

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: any n or Xaa = unknown

US-09-377-497-70

Query Match

Best Local Similarity

Matches 7; Conservative

35.5%; Score 39; DB 4; Length 664;

Pred. No. 2.7e+02;

Mismatches 2; Indels 0; Gaps 0;

QY 1 QARKKTFLSVHEV 13

DB 417 RARSKTFINLREV 429

RESULT 15

US-09-422-869-28

; Sequence 28, Application US/09422869

; Patent No. 6235481

; GENERAL INFORMATION:

; APPLICANT: POLONSKY, KENNETH S.

; APPLICANT: HORIKAWA, YUKIO

; APPLICANT: ODA, NAOHISA

; APPLICANT: COX, NANCY J.

; APPLICANT: SEENAN, SEAMUS

; APPLICANT: ZHOU, YUN-PING

; APPLICANT: OTANI, KENICHI

; APPLICANT: HANIS, CRAIG L.

; APPLICANT: BELL, GRAEME I.

; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES

; FILE REFERENCE: ARCD:307

; CURRENT APPLICATION NUMBER: US/09/422,869

; CURRENT FILING DATE: 1999-10-21

; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27435

Query Match 36.8%; Score 40.5; DB 4; Length 293;  
Best Local Similarity 36.8%; Pred. No. 61;  
Matches 7; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 4 KKTFLSVHEVMAVENYAKD 22  
Db 207 RKHFID-HEINSIQNYLSD 224

RESULT 7

US-09-800-729-195  
; Sequence 195, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; FILE OF INVENTION: 32 Human secreted proteins  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 195  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-195

Query Match 36.4%; Score 40; DB 4; Length 326;  
Best Local Similarity 31.2%; Pred. No. 84;  
Matches 5; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 4 KKTFLSVHEVMAVENY 19  
Db 8 KKIFIHMHBIQIDGH 23

RESULT 8

US-09-489-039A-13335  
; Sequence 1335, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 13335  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-13335

Query Match 36.4%; Score 40; DB 4; Length 419;  
Best Local Similarity 47.4%; Pred. No. 1.1e+02;  
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 4 KKTFLSVHEVMAVENYAKD 22  
Db 253 KATFSAVHEAQAYETRPVD 271

RESULT 9

US-09-252-991A-25843  
; Sequence 25843, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25843  
; LENGTH: 1242  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25843

Query Match 36.4%; Score 40; DB 4; Length 1242;  
Best Local Similarity 50.0%; Pred. No. 3.7e+02;  
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 5 KTFSLVHEVMAVENYAKD 22  
Db 320 ETFGVIEETTAVERLAKD 337

RESULT 10

US-09-540-236-2056  
; Sequence 2056, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRI  
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 2056  
; LENGTH: 605  
; TYPE: PRT  
; ORGANISM: M. catarrhalis  
US-09-540-236-2056

Query Match 35.9%; Score 39.5; DB 4; Length 605;  
Best Local Similarity 32.4%; Pred. No. 2e+02;  
Matches 11; Conservative 2; Mismatches 2; Indels 19; Gaps 1;

QY 8 LSVHEVMAVE-----NYAKD 22  
Db 89 LSQHQVMAIEALAVYVMAVLFVGMALLRALLIGYAKD 122

RESULT 11

US-09-328-352-6399  
; Sequence 6399, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6399  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii

Matches	10;	Conservative	5;	Mismatches	6;	Indels	2;	Gaps	1;
QY	1	QARKKTFLSVHEVM--AVENYAK 21							
					:				
Db	227	RARHKAGLCHDILLPESVINYK 249							
					:				
RESULT 3									
US-09-107-532A-6220									
;		Sequence 6220, Application US/09107532A							
;		Patent No. 6583275							
;		GENERAL INFORMATION:							
;		APPLICANT: Lynn A Doucette-Stamm and David Bush							
;		TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO							
;		ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS							
;		NUMBER OF SEQUENCES: 7310							
;		CORRESPONDENCE ADDRESS:							
;		ADDRESSEE: GENOME THERAPEUTICS CORPORATION							
;		STREET: 100 Beaver Street							
;		CITY: Waltham							
;		STATE: Massachusetts							
;		COUNTRY: USA							
;		ZIP: 02354							
;		COMPUTER READABLE FORM:							
;		MEDIUM TYPE: CD-ROM ISO9660							
;		COMPUTER: PC							
;		OPERATING SYSTEM: <Unknown>							
;		SOFTWARE: ASCII							
;		CURRENT APPLICATION DATA:							
;		APPLICATION NUMBER: US/09/107,532A							
;		FILING DATE: 30-Jun-1998							
;		PRIOR APPLICATION DATA:							
;		APPLICATION NUMBER: 60/085,598							
;		FILING DATE: 14 May 1998							
;		APPLICATION NUMBER: 60/051571							
;		FILING DATE: July 2, 1997							
;		ATTORNEY/AGENT INFORMATION:							
;		NAME: Ariniello, Pamela Deneke							
;		REGISTRATION NUMBER: 40,489							
;		REFERENCE/DOCKET NUMBER: GTC-012							
;		TELECOMMUNICATION INFORMATION:							
;		TELEPHONE: (781)893-5007							
;		TELEFAX: (781)893-8277							
;		INFORMATION FOR SEQ ID NO: 6220:							
;		SEQUENCE CHARACTERISTICS:							
;		LENGTH: 1224 amino acids							
;		TYPE: amino acid							
;		TOPOLOGY: linear							
;		MOLECULE TYPE: protein							
;		HYPOTHETICAL: YES							
;		ORIGINAL SOURCE:							
;		ORGANISM: Enterococcus faecium							
;		FEATURE:							
;		NAME/KEY: misc_feature							
;		LOCATION: (B) LOCATION 1...1224							
;		SEQUENCE DESCRIPTION: SEQ ID NO: 6220:							
US-09-107-532A-6220									
Query Match		39.1%;	Score 43;	DB 4;	Length 1224;				
Best Local Similarity		53.3%;	Pred. No. 1.2e+02;						
Matches	8;	Conservative	3;	Mismatches	4;	Indels	0;	Gaps	0;
QY	5	KTFLSVHEVMVAVENY 19							
Db	1040	KQLQVRDVLVSNY 1054							
RESULT 4									
US-09-252-991A-30881									
;		Sequence 30881, Application US/09252991A							
;		Patent No. 6551795							
;		GENERAL INFORMATION:							
;		APPLICANT: Marc J. Rubenfield et al.							

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 5.63095 Seconds  
(without alignments)  
201.701 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKTFLSVHEVAVENYAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pcp.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	44.5	195	4	US-09-198-452A-616
2	43	39.1	255	4	US-09-543-681A-5383
3	43	39.1	1224	4	US-09-107-532A-6220
4	41	37.3	316	4	US-09-252-991A-30881
5	41	37.3	839	4	US-09-489-039A-13252
6	40.5	36.8	293	4	US-09-252-991A-27435
7	40	36.4	326	4	US-09-800-729-195
8	40	36.4	419	4	US-09-489-039A-13335
9	40	36.4	1242	4	US-09-252-991A-25843
10	39.5	35.9	605	4	US-09-540-236-2056
11	39	35.5	291	4	US-09-328-352-6399
12	39	35.5	399	4	US-09-134-000C-5712
13	39	35.5	469	3	US-08-378-313-33
14	39	35.5	664	4	US-09-377-497-70
15	39	35.5	690	3	US-09-422-869-28
16	39	35.5	1284	4	US-09-170-496B-294
17	39	35.5	1284	4	US-09-364-425B-59
18	39	35.5	1507	3	US-08-929-329-5
19	38	34.5	133	1	US-08-554-463-1
20	38	34.5	208	4	US-09-540-236-3351
21	38	34.5	245	4	US-09-489-039A-12667
22	38	34.5	271	2	US-08-956-012-3
23	38	34.5	359	4	US-09-543-681A-5431
24	38	34.5	434	4	US-09-543-681A-4577
25	38	34.5	501	1	US-08-451-715A-10
26	38	34.5	611	4	US-09-107-532A-4988
27	38	34.5	715	4	US-09-620-412C-329

Sequence 329, App  
Sequence 29, Appl  
Sequence 29, Appl  
Sequence 30, Appl  
Sequence 30, Appl  
Sequence 3159, Ap  
Sequence 29131, A  
Sequence 4, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 23, Appl  
Sequence 22, Appl  
Sequence 6, Appli  
Sequence 3060, Ap  
Sequence 85, Appl  
Sequence 90, Appl  
Sequence 4234, Ap  
Sequence 7778, Ap

28 38 34.5 715 4 US-09-598-419-329  
29 38 34.5 2987 2 US-08-970-269A-29  
30 38 34.5 2987 3 US-09-407-562-29  
31 38 34.5 3959 2 US-08-970-269A-30  
32 38 34.5 3959 3 US-09-407-562-30  
33 38 34.5 10182 4 US-09-134-001C-3159  
34 37.5 34.1 336 4 US-09-252-991A-29131  
35 37.5 34.1 578 1 US-08-766-014-4  
36 37.5 34.1 608 1 US-08-766-014-3  
37 37.5 34.1 872 1 US-08-766-014-2  
38 37 33.6 38 4 US-09-779-451-23  
39 37 33.6 45 4 US-09-779-451-22  
40 37 33.6 138 4 US-09-570-921-6  
41 37 33.6 190 4 US-09-540-236-3060  
42 37 33.6 198 3 US-08-965-056-85  
43 37 33.6 198 3 US-08-965-056-90  
44 37 33.6 207 4 US-09-543-681A-4234  
45 37 33.6 221 4 US-09-489-039A-7778

#### ALIGNMENTS

##### RESULT 1

US-09-198-452A-616  
; Sequence 616, Application US/09198452A  
; Patent No. 6559294

; GENERAL INFORMATION:

; APPLICANT: Griffiths, R.

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments

; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention

; FILE REFERENCE: 9710-003-999

; CURRENT APPLICATION NUMBER: US/09/198,452A

; CURRENT FILING DATE: 1998-11-24

; NUMBER OF SEQ ID NOS: 6849

; SEQ ID NO 616

; LENGTH: 195

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-198-452A-616

Query Match 44.5%; Score 49; DB 4; Length 195;  
Best Local Similarity 50.0%; Pred. No. 1.6;  
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVAVENYAKD 22

Db 37 RQKTFLLSHRVLARPSFPVD 56

##### RESULT 2

US-09-543-681A-5383  
; Sequence 5383, Application US/09543681A  
; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 5383

; LENGTH: 255

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-5383

Query Match 39.1%; Score 43; DB 4; Length 255;  
Best Local Similarity 43.5%; Pred. No. 20;

XX 09-OCT-2001 (first entry)  
 DT Peptide #1515 encoded by probe for measuring breast gene expression.  
 XX  
 DE Probe; human; breast disease; breast cancer; development disorder;  
 XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 KW  
 KW Homo sapiens.  
 XX  
 OS WO200157270-A2.  
 XX  
 PN 09-AUG-2001.  
 XX  
 PD 29-JAN-2001; 2001WO-US000661.  
 XX  
 PF 04-FEB-2000; 2000US-0180312P.  
 XX  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR  
 PR 30-JUN-2000; 2000US-00608408.  
 PR  
 PR 03-AUG-2000; 2000US-00632366.  
 PR  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-476286/51.  
 XX  
 DR Novel single exon nucleic acid probe used to measuring gene expression in  
 XX a human breast.  
 PT  
 PT Claim 27; SEQ ID NO 11573; 322pp; English.  
 XX  
 PS The present invention relates to novel single exon nucleic acid probes  
 XX (see AA100010-AA110067). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for measuring human gene expression in  
 CC a human breast sample, where the probe hybridises at high stringency to a  
 CC nucleic acid expressed in the human breast. The probes are useful for  
 CC predicting, diagnosing, grading, staging, monitoring and prognosing  
 CC diseases of the human breast, particularly those diseases with polygenic  
 CC aetiology. The diseases include: breast cancer, disorders of development,  
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
 CC breast disease and non-carcinoma tumours. Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 50 AA;  
 SQ  
 Query Match 91.8%; Score 101; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 RKKTFLSVHEVMAVENYAKD 22  
 Db 1 RKKTFLSVHEVMAVENYAKD 20

Search completed: March 18, 2004, 14:14:57  
 Job time : 20.9484 secs

CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention

XX Sequence 50 AA;

Query Match 91.8%; Score 101; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFSLSVHEVMAVENYAKD 22  
 |||||  
 Db 1 RKKTFSLSVHEVMAVENYAKD 20

RESULT 13  
 AAM54871  
 ID AAM54871 standard; protein; 50 AA.

XX AC AAM54871;

DT 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 26976.

DE Human; brain expressed exon; gene expression analysis; probe; microarray;

XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000667.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX brains.

XX Example 4; SEQ ID NO 26976; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX brain. They can be used to measure gene expression in brain cell samples,

XX which may enable the diagnosis and improved treatment of nervous system

XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX epilepsy and cancers. The present sequence is a protein encoded by one of

XX the probes of the invention

RESULT 14

ABG48915

ID ABG48915 standard; peptide; 50 AA.

XX AC ABG48915;

DT 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 27563.

DE Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

XX hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000664.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human adult liver.

XX Claim 27; SEQ ID NO 27563; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for

XX measuring human gene expression in a sample derived from human adult

XX liver, comprising one of 13109 defined nucleotide sequences given in the

XX specification for complements/ fragments). The probe hybridises at high

XX stringency to a nucleic acid molecule expressed in the human adult liver.

XX (I) may be used for predicting, measuring and displaying gene expression

XX in samples derived from human adult liver. The genes identified may be

XX involved in genetic liver diseases such as cirrhosis,

XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is

XX associated with coronary heart disease. ABG47348-ABG59930 represent human

XX liver single exon encoded peptides of the invention. Note: The sequence

XX information for this patent does not appear in the printed specification

XX but was obtained in electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 50 AA;

Query Match 91.8%; Score 101; DB 4; Length 50;

Best Local Similarity 100.0%; Pred. No. 1.5e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFSLSVHEVMAVENYAKD 22

|||||  
 Db 1 RKKTFSLSVHEVMAVENYAKD 20

RESULT 15

AAM02833

ID AAM02833 standard; protein; 50 AA.

XX AC AAM02833;

XX WPI; 2001-496933/54.  
 XX New spatially-addressable set of single exon nucleic acid probes, useful  
 PT for measuring gene expression in sample derived from human breast,  
 PT comprises number of single exon nucleic acid probes.  
 XX  
 PS Claim 27; SEQ ID NO 11881; 327pp + Sequence Listing; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and B7 474 cells. The method involves contacting the  
 CC probes with a collection of detectably labelled nucleic acids derived  
 CC from mRNA of human breast, and then measuring the label bound to each  
 CC probe of the microarray. The probes are useful for verifying the  
 CC expression of regions of genomic DNA predicted to encode proteins. They  
 CC are useful for gene discovery, and for determining predisposition and/or  
 CC prognosing breast disease. Gene expression analysis is useful for  
 CC assessing the toxicity of chemical agents on cells. The microarray of  
 CC this invention presents a far greater diversity of probes for measuring  
 CC gene expression, with far less bias than expressed sequence tag  
 CC microarrays. The method is suitable for rapid production of functional  
 CC information from genomic sequence. The present sequence is a peptide  
 CC encoded by a single exon nucleic acid probe of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 50 AA;  
 Query Match 91.8%; Score 101; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RKKTFSLVHEVMAVENYAKD 22  
 DB 1 RKKTFSLVHEVMAVENYAKD 20  
 RESULT 11  
 ABB19524  
 ID ABB19524 standard; protein; 50 AA.  
 AC ABB19524;  
 DT 23-JAN-2002 (first entry)  
 DE Protein #1523 encoded by probe for measuring heart cell gene expression.  
 DE Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 XX Homo sapiens.  
 OS WO200157274-A2.  
 PN 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US000666.  
 PF 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488900/53.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.  
 XX Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human

DR WPI; 2001-488999/53.  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts.  
 PS Claim 15; SEQ ID NO 21294; 530pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA411305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 50 AA;  
 Query Match 91.8%; Score 101; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RKKTFSLVHEVMAVENYAKD 22  
 DB 1 RKKTFSLVHEVMAVENYAKD 20  
 RESULT 12  
 AAM67252  
 ID AAM67252 standard; protein; 50 AA.  
 AC AAM67252;  
 DT 06-NOV-2001 (first entry)  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27558.  
 DE Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX Homo sapiens.  
 OS WO200157276-A2.  
 PN 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US000668.  
 PF 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488900/53.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.  
 XX Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human

PN WO200157272-A2.  
 XX 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US0000663.  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488897/53.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.  
 XX  
 XX Claim 27; SEQ ID NO 27814; 654pp; English.  
 PS The present invention relates to single exon nucleic acid probes (SENP:  
 CC see AAI31315-AA157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders  
 XX human genetic disorders  
 XX Sequence 50 AA;  
 SQ  
 Query Match 91.8%; Score 101; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RKKTFLSVHEVMAVENYAKD 22  
 Db 1 RKKTFLSVHEVMAVENYAKD 20  
 RESULT 9  
 ABB23289  
 ID ABB23289 standard; peptide; 50 AA.  
 XX  
 AC ABB23289;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Peptide #5040 encoded by breast cell single exon nucleic acid probe.  
 XX Human; microarray; single exon probe; gene expression; breast; disease;  
 KW cancer.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157271-A2.  
 XX  
 XX 09-AUG-2001.  
 XX  
 XX 30-JAN-2001; 2001WO-US0000662.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-496933/54.  
 XX  
 XX New spatially-addressable set of single exon nucleic acid probes, useful  
 PT for measuring gene expression in sample derived from human breast,  
 PT comprises number of single exon nucleic acid probes.  
 XX  
 XX Claim 27; SEQ ID NO 15357; 327pp + Sequence Listing; English.  
 XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting the  
 CC probes with a collection of detectably labelled nucleic acids derived  
 CC from mRNA of human breast, and then measuring the label bound to each  
 CC probe of the microarray. The probes are useful for verifying the  
 CC expression of regions of genomic DNA predicted to encode proteins. They  
 CC are useful for gene discovery, and for determining predisposition and/or  
 CC assessing the toxicity of chemical agents on cells. The microarray of  
 CC this invention presents a far greater diversity of probes for measuring  
 CC gene expression, with far less bias than expressed sequence tag  
 CC microarrays. The method is suitable for rapid production of functional  
 CC information from genomic sequence. The present sequence is a peptide  
 CC encoded by a single exon nucleic acid probe of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 50 AA;  
 SQ  
 Query Match 91.8%; Score 101; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RKKTFLSVHEVMAVENYAKD 22  
 Db 1 RKKTFLSVHEVMAVENYAKD 20  
 RESULT 10  
 ABB28913  
 ID ABB28913 standard; peptide; 50 AA.  
 XX  
 AC ABB28913;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Peptide #1564 encoded by breast cell single exon nucleic acid probe.  
 XX Human; microarray; single exon probe; gene expression; breast; disease;  
 KW cancer.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157271-A2.  
 XX  
 XX 09-AUG-2001.  
 XX  
 XX 30-JAN-2001; 2001WO-US0000662.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI

Query Match 91.8%; Score 101; DB 5; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 1e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHVMVAVENYAKD 22  
 |||||  
 Db 1 RKKTFLSVHVMVAVENYAKD 20

RESULT 6  
 AM15096  
 ID AAM15096 standard; protein; 50 AA.  
 XX AC AAM15096;  
 XX DT 12-OCT-2001 (first entry)  
 XX DE Peptide #1530 encoded by probe for measuring cervical gene expression.  
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 XX KW cervical cancer.  
 XX OS Homo sapiens.  
 XX PN WO200157278-A2.  
 XX PD 09-AUG-2001.  
 XX PF 30-JAN-2001; 2001WO-US0000670.  
 XX PR 04-FEB-2000; 2000US-0180312P.  
 XX PR 26-MAY-2000; 2000US-0207456P.  
 XX PR 30-JUN-2000; 2000US-00608408.  
 XX PR 03-AUG-2000; 2000US-00632366.  
 XX PR 21-SEP-2000; 2000US-0234687P.  
 XX PR 27-SEP-2000; 2000US-0236359P.  
 XX PR 04-OCT-2000; 2000GB-00024263.  
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX PF WPI; 2001-488901/53.  
 XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 XX PT gene expression in human cervical epithelial cells.  
 XX PS Claim 27; SEQ ID NO 19922; 487pp; English.  
 XX CC The present invention relates to human single exon nucleic acid probes  
 CC (SENPs; see AAI10068-AA128459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 50 AA;

Query Match 91.8%; Score 101; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHVMVAVENYAKD 22  
 |||||  
 Db 1 RKKTFLSVHVMVAVENYAKD 20

RESULT 7

ABB34086  
 ID ABB34086 standard; peptide; 50 AA.  
 XX AC ABB34086;  
 XX DT 04-FEB-2002 (first entry)  
 XX DE Peptide #1592 encoded by human foetal liver single exon probe.  
 XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX OS Homo sapiens.  
 XX PN WO200157277-A2.  
 XX PD 09-AUG-2001.  
 XX PF 30-JAN-2001; 2001WO-US0000669.  
 XX PR 04-FEB-2000; 2000US-0180312P.  
 XX PR 26-MAY-2000; 2000US-0207456P.  
 XX PR 30-JUN-2000; 2000US-00608408.  
 XX PR 03-AUG-2000; 2000US-00632366.  
 XX PR 21-SEP-2000; 2000US-0234687P.  
 XX PR 27-SEP-2000; 2000US-0236359P.  
 XX PR 04-OCT-2000; 2000GB-00024263.  
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX PF WPI; 2001-483447/52.  
 XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 XX PT gene expression in human fetal liver.  
 XX PS Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.  
 XX CC The invention relates to a single exon nucleic acid probe for measuring  
 CC human gene expression in a sample derived from human foetal liver. The  
 CC single exon nucleic acid probes may be used for predicting, measuring and  
 CC displaying gene expression in samples derived from human fetal liver. The  
 CC present sequence is a peptide encoded by a single exon nucleic acid probe  
 CC of the invention. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 50 AA;  
 XX Query Match 91.8%; Score 101; DB 4; Length 50;  
 XX Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
 XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHVMVAVENYAKD 22  
 |||||  
 Db 1 RKKTFLSVHVMVAVENYAKD 20

RESULT 8  
 AM27545  
 ID AAM27545 standard; protein; 50 AA.  
 XX AC AM27545;  
 XX DT 17-OCT-2001 (first entry)  
 XX DE Peptide #1582 encoded by probe for measuring placental gene expression.  
 XX KW Probe; microarray; human; placenta; antenatal diagnosis;  
 XX KW genetic disorder.  
 XX OS Homo sapiens.  
 XX

CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 XX  
 SQ Sequence 117 AA;  
 Query Match 100.0%; Score 110; DB 5; Length 117;  
 Best Local Similarity 100.0%; Pred. NO. 1.4e-09;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QARKKTFSLVHEVMAVENYAKD 22  
 DB 4 QARKKTFSLVHEVMAVENYAKD 25  
 RESULT 4  
 AAU79852  
 ID AAU79852 standard; protein; 137 AA.  
 XX  
 AC AAU79852;  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8).  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 DR N-PSDB; ABK49522.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 93-94; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This is the amino acid sequence of human cystatin-8 (Zcys8)  
 XX  
 SQ Sequence 137 AA;  
 Query Match 100.0%; Score 110; DB 5; Length 137;  
 Best Local Similarity 100.0%; Pred. NO. 1.7e-09;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QARKKTFSLVHEVMAVENYAKD 22  
 DB 24 QARKKTFSLVHEVMAVENYAKD 45  
 RESULT 5  
 AAU79856  
 ID AAU79856 standard; peptide; 36 AA.  
 XX  
 AC AAU79856;  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #4.  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic fragment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 DR Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 DR spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 DR to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 95; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 XX  
 SQ Sequence 36 AA;

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 CC  
 XX  
 SQ Sequence 22 AA;  
 Query Match 100.0%; Score 110; DB 5; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-10;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QARKKTFSLVHEVMAVENYAKD 22  
 |||||  
 DB 1 QARKKTFSLVHEVMAVENYAKD 22  
 |||||  
 RESULT 2  
 AAU79853  
 ID AAU79853 standard; protein; 115 AA.  
 XX  
 AC AAU79853;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #1.  
 XX  
 DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic fragment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026968.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 94; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8

CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 CC  
 XX  
 SQ Sequence 115 AA;  
 Query Match 100.0%; Score 110; DB 5; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QARKKTFSLVHEVMAVENYAKD 22  
 |||||  
 DB 2 QARKKTFSLVHEVMAVENYAKD 23  
 |||||  
 RESULT 3  
 AAU79854  
 ID AAU79854 standard; protein; 117 AA.  
 XX  
 AC AAU79854;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #2.  
 XX  
 DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic fragment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 94-95; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 19,9484 Seconds  
(without alignments)  
311.606 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKKTFSLVHEWVAENVAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	110	100.0	22	5 AAU79855	Aau79855 Human cys
2	110	100.0	115	5 AAU79853	Aau79853 Human cys
3	110	100.0	117	5 AAU79854	Aau79854 Human cys
4	110	100.0	137	5 AAU79852	Aau79852 Human cys
5	101	91.8	36	5 AAU79856	Aau79856 Human cys
6	101	91.8	50	4 AAM15096	Aam15096 Peptide #
7	101	91.8	50	4 ABB34086	Abb34086 Peptide #
8	101	91.8	50	4 AAM27545	Aam27545 Peptide #
9	101	91.8	50	4 ABB32389	Abb32389 Peptide #
10	101	91.8	50	4 ABB28913	Abb28913 Peptide #
11	101	91.8	50	4 ABB19524	Abb19524 Protein #
12	101	91.8	50	4 AAM67252	Aam67252 Human bon
13	101	91.8	50	4 AAM54871	Aam54871 Human bra
14	101	91.8	50	4 ABB48915	Abb48915 Human liv
15	101	91.8	50	4 AAM02833	Aam02833 Peptide #
16	101	91.8	50	5 ABB36903	Abb36903 Human pep
17	53	48.2	184	5 ABB30713	Abb30713 Chlamydia
18	49	44.5	195	2 AAY35198	Aay35198 Chlamydia
19	45	40.9	517	4 ABB57957	Abb57957 Drosophil
20	44	40.0	180	6 ADB09534	Adb09534 Allostoc
21	44	40.0	296	4 AAM07104	Aam07104 Human nov
22	44	40.0	318	4 AAU07106	Aau07106 Human nov
23	44	40.0	370	5 ABB97599	Abb97599 Novel hum
24	44	40.0	375	4 AAU07109	Aau07109 Human nov
25	44	40.0	473	4 AAU07111	Aau07111 Human nov

## ALIGNMENTS

### RESULT 1

AAU79855  
ID AAU79855 standard; peptide; 22 AA.

XX AAU79855;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #3.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic fragment.

XX Homo sapiens.

XX WO200220567-A2.

PD 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO ) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 95; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8) polypeptide (I). (i) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcys8 is useful as research reagent for characterising sites of interaction between Zcys8 and its receptor. Zcys8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(i) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcys8. The antibodies are also useful to isolate large

Adc51944 Human pos  
Aau07115 Human nov  
Ada55525 Human pro  
Adc10176 Human NOV  
Aau07119 Human nov  
Aau07123 Human kin  
Aae16270 Human mem  
Abb79941 Human pos  
Adc51943 Human pos  
Aau20555 Human sec  
Aaw98554 H. pylori  
Abu37832 Protein e  
Ades2552 Caenorhab  
Aau33499 Enterococ  
Abu29638 Protein e  
Aau35034 Enterococ  
Abu14481 Protein e  
Adc96593 E. faeciu  
Abu33247 Protein e  
Abb17531 Human ner

26 44 40.0 491 7 ADC51944  
27 44 40.0 520 4 AAU07115  
28 44 40.0 542 6 ADA55525  
29 44 40.0 549 7 ADC10176  
30 44 40.0 560 4 AAU07119  
31 44 40.0 576 4 AAU07123  
32 44 40.0 576 5 AAE16270  
33 44 40.0 576 5 ABB79941  
34 44 40.0 576 7 ADC51943  
35 44 40.0 580 4 AAU20555  
36 43 39.1 259 2 AAW98554  
37 43 39.1 431 6 ABU37832  
38 43 39.1 467 7 ADE52552  
39 43 39.1 1054 4 AAU33499  
40 43 39.1 1175 6 ABU29638  
41 43 39.1 1208 4 AAU35034  
42 43 39.1 1217 6 ABU14481  
43 43 39.1 1224 7 ADC96593  
44 42.5 38.6 281 6 ABU33247  
45 42 38.2 139 4 ABB17531

Search completed: March 18, 2004, 14:21:23  
Job time : 70.1071 secs

```
RA Ikegawa S., Nakamura Y.;
RT "DD72, a novel mouse gene implicated in the early stage of ectopic
RL ossification.";
RN Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036743; BAA95411.1; -
DR EMBL; BC048364; AAH48364.1; -
DR HSSP; P01034; IG96.
DR MGD; MGI:1930004; Cst10.
DR GO; GO:0004869; F:Cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR001713; Stefina.
DR Pfam; PF00031; cystatin; 1.
DR PRINTS; PR00295; STEPINA.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
SQ SEQUENCE 148 AA; 16451 MW; 637534CBFCSAA179 CRC64;

Query Match 17.7%; Score 113; DB 11; Length 148;
Best Local Similarity 28.2%; Pred. No. 0.00031;
Matches 29; Conservative 18; Mismatches 46; Indels 10; Gaps 3;

QY 15 EVMAVENYAKDSLQWITDQYNKESDDKYHPRIFRVLKQVQVTDHLEYHLNVEMQWTTCC 74
DB 50 EVQKVKFA-----VRTYNDMDNDLYLSKPIRLMSASQVVAGKNYYLKIELGRTTCT 102

QY 75 KPET--TNCVQPQR-ELHKQVNCFFSVFVFPWFQYKILNKSC 114
DB 103 KTESNLVDCPFNEQPDQKRVICNFQINVAFLNKNKSMTNFNC 145

RESULT 15
Q8K397
ID Q8K397 PRELIMINARY; PRT; 146 AA.
AC Q8K397;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE RIKEN cDNA 110017E11 gene (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027680; AAH27680.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
FT NON TER 1
FT SEQUENCE 146 AA; 16380 MW; 9D77BB9A6063A5C4 CRC64;

Query Match 17.6%; Score 112.5; DB 11; Length 146;
Best Local Similarity 26.1%; Pred. No. 0.00034;
Matches 23; Conservative 20; Mismatches 38; Indels 7; Gaps 1;

QY 34 YNKESDDKYHPRIFRVLKQVQVTDHLEYHLNVEMQWTTCKP-----ETTNCVQPQR 86
DB 56 YNMGSDSLYYFRDTKVIDAKYQLVAGIKYLLDIESTECRKTQVSGEHMDLTTCPLAAG 115

QY 87 ELHKQVNCFFSVFVFPWFQYKILNKSC 114
DB 116 GQEKLRNCFNLELVFPWNTTQLKHDC 143
```

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AK020193; BAB32024.1; -

DR HSSP; P01034; IG96.

DR MGD; MGI:1925859; 8030411P24rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; Cy; 1

DR SEQUENCE 130 AA; 14947 MW; DD2F930B64B4E584 CRC64;

Query Match 18.9%; Score 121; DB 11; Length 130;

Best Local Similarity 27.5%; Pred. No. 3.8e-05;

Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 8 KTFSLVHEVMAVENYAKDSQWITDQYNKESDDKHFRIFRLVKVQRQVTDHLYHLNVE 67

Db 23 KEFLDVTKDL---DYFVASVEFAVAQFNNDNNPEENTYKLLLEVGRQAKK-TWTMIFLMDLE 78

QY 68 MOWTTCQKPETT--NCVPQRELHKQVNCFFSFAVPWFQYKILNKSC 114

Db 79 MGRITCKKHNDHINCPQLQOSREKKVHCVFQVDARPFWSHFILTSTC 127

RESULT 12

Q8VLI3 PRELIMINARY; PRT; 130 AA.

AC Q8VLI3

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Cystatin SC.

GN 8030411P24rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57; TISSUE=Testis;

RA Li Y., Friel P.J., Griswold M.D.;

RT "Molecular cloning and characterization of cystatin SC and cystatin

TE-1, new members of the cystatin family.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF440735; AAL30841.1; -

DR MGD; MGI:1925859; 8030411P24rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; Cy; 1

DR SEQUENCE 130 AA; 15076 MW; DD34930B64AFES9F CRC64;

Query Match 18.9%; Score 121; DB 11; Length 130;

Best Local Similarity 27.5%; Pred. No. 3.8e-05;

Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 8 KTFSLVHEVMAVENYAKDSQWITDQYNKESDDKHFRIFRLVKVQRQVTDHLYHLNVE 67

Db 23 KEFLDVTKDL---DYFVASVEFAVAQFNNDNNPEENTYKLLLEVGRQAKK-TWTMIFLMDLE 78

QY 68 MOWTTCQKPETT--NCVPQRELHKQVNCFFSFAVPWFQYKILNKSC 114

Db 79 MGRITCKKHNDHINCPQLQOSREKKVHCVFQVDARPFWSHFILTSTC 127

RESULT 13

Q9DAN8 PRELIMINARY; PRT; 128 AA.

AC Q9DAN8

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE 1700006F03rik protein (Cystatin TE-1).

GN 1700006F03rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bargh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wyshaw-Borja A., Yoshida K., Hasegawa Y., Kawaji H., Kohtauki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57; TISSUE=Testis;

RA Li Y., Friel P.J., Griswold M.D.;

RT "Molecular cloning and characterization of cystatin SC and cystatin

TE-1, new members of the cystatin family.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK005670; BAB24179.1; -

DR EMBL; AF440737; AAL30843.1; -

DR MGD; MGI:1916612; 1700006F03rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; Cy; 1.

DR SEQUENCE 128 AA; 15036 MW; 40AFD00103E5ED6 CRC64;

Query Match 17.9%; Score 114.5; DB 11; Length 128;

Best Local Similarity 29.7%; Pred. No. 0.00018;

Matches 27; Conservative 20; Mismatches 41; Indels 3; Gaps 2;

QY 26 SLQWITDQYNKESDDKHFRIFRLVKVQRQVTDHLYHLNVEQWTTCCQ--PETTNCVP 83

Db 36 SVEHVVFHFNENQDDDFAYKFLVRRLRSLQKYT-LKYLVDLEMGRTLCGKYDEIDNCPL 94

QY 84 QRELHKQVNCFFSFAVPWFQYKILNKSC 114

Db 95 QEGPGRKVRCTIVTEAWTKFTILNSTC 125

RESULT 14

Q9JMB4 PRELIMINARY; PRT; 148 AA.

ID Q9JMB4

AC Q9JMB4;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE DD72 protein (similar to cystatin 10) (Chondrocytes).

GN CST10 OR DD72.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RA  Bai J., Lao H., Ye X., Li Y., Lou J.;
RT  "Molecular cloning and sequence analysis of cystatin cDNA from two
RL  species of sturgeons.";
DR  Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF334610; AAK16731.1; -.
DR  HSSP; P01038; 1A90.
DR  GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR  InterPro; IPR000010; Cystatin.
DR  Pfam; PF00031; cystatin; 1.
DR  SMART; SM00043; CY; 1.
DR  PROSITE; PS00287; CYSTATIN; 1.
FT  NON TER
SQ  SEQUENCE 112 AA; 12231 MW; 48ECEBFD8A08C00 CRC64;

Query Match 21.7%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 3.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY 20 ENYAKDSLOWITDQYNKESDDKYHFRFVLKQVQRTDHLVHVMQWTTTCQK---P 76
Db 13 EEGVQDALKFAVAEFNKASNDMTIHRVSKVKQVQVAGIKYIVTQMGRTSCKRGGA 72

QY 77 ETTNC-----VPOBELHKLQVNCFFSVPFVAPWPFQYKILNKSCS 115
Db 73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLKVNKTCT 112

RESULT 9
Q98SR3 PRELIMINARY; PRT; 112 AA.
AC Q98SR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin (Fragment).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=111304;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bai J., Lao H., Ye X., Li Y., Lou J.;
RT "Molecular cloning and sequence analysis of cystatin cDNA from two
RL species of sturgeons.";
DR Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF334611; AAK16732.1; -.
DR HSSP; P01038; 1A90.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT NON TER
SQ SEQUENCE 112 AA; 12231 MW; 48ECEBFD8A08C00 CRC64;

Query Match 21.7%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 3.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY 20 ENYAKDSLOWITDQYNKESDDKYHFRFVLKQVQRTDHLVHVMQWTTTCQK---P 76
Db 13 EEGVQDALKFAVAEFNKASNDMTIHRVSKVKQVQVAGIKYIVTQMGRTSCKRGGA 72

QY 77 ETTNC-----VPOBELHKLQVNCFFSVPFVAPWPFQYKILNKSCS 115
Db 73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLKVNKTCT 112

RESULT 10
Q98SR3 PRELIMINARY; PRT; 112 AA.
AC Q98SR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin (Fragment).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=111304;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bai J., Lao H., Ye X., Li Y., Lou J.;
RT "Molecular cloning and sequence analysis of cystatin cDNA from two
RL species of sturgeons.";
DR Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF334611; AAK16732.1; -.
DR HSSP; P01038; 1A90.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT NON TER
SQ SEQUENCE 112 AA; 12231 MW; 48ECEBFD8A08C00 CRC64;

Query Match 21.7%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 3.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY 20 ENYAKDSLOWITDQYNKESDDKYHFRFVLKQVQRTDHLVHVMQWTTTCQK---P 76
Db 13 EEGVQDALKFAVAEFNKASNDMTIHRVSKVKQVQVAGIKYIVTQMGRTSCKRGGA 72

QY 77 ETTNC-----VPOBELHKLQVNCFFSVPFVAPWPFQYKILNKSCS 115
Db 73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLKVNKTCT 112
```

```

Q8VIH8 PRELIMINARY; PRT; 130 AA.
AC Q8VIH8;
DT 01-WAR-2002 (TrEMBLrel. 20, Created)
DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin SC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RA Li Y., Friel P.J., Griswold M.D.;
RT "Molecular cloning and characterization of cystatin SC and cystatin
RL TE-1, new members of the cystatin family.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF442205; AAL35350.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 130 AA; 14981 MW; 7A752359860989C9 CRC64;

Query Match 19.1%; Score 122; DB 11; Length 130;
Best Local Similarity 27.5%; Pred. No. 3e-05;
Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 8 KTFSLVHVMVAVENYAKDSLOWITDQYNKESDDKYHFRFVLKQVQRTDHLVHVMQWTTTCQK---P 76
Db 23 KEFLDVTKDL---DYFVASVFVAQFNDNNSEENTYRLLEVGRAQKK-TWTMFLMDLE 78

QY 68 MOWTTTCQKPETT--NCVPOBELHKLQVNCFFSVPFVAPWPFQYKILNKSC 114
Db 79 MGRITCKKHDEHNCPLLOGSGEKKVCHVQVDARPFWSHFVTLSTC 127

RESULT 11
Q9CX46 PRELIMINARY; PRT; 130 AA.
AC Q9CX46;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 8030411P24RIK protein.
GN 8030411P24RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic testis;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Garioldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,
RA Hayaishizaki Y.,
```

```
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048681; AAH48681.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR DR Pfam; PF00031; cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 141 AA; 16825 MW; C20FA0D8A894951F CRC64;

Query Match 29.5%; Score 188.5; DB 11; Length 141;
Best Local Similarity 34.5%; Pred. No. 2.8e-12;
Matches 40; Conservative 25; Mismatches 38; Indels 13; Gaps 3;

QY 2 PYQARKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFVLKVRQVTDHLE 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
33 PFEDIPKSYV-----YVQHALWYAMKEYNKASNDLYNFRVVDILKSQEQITDSLE 82

QY 62 YHLNVEQWTTCK--PETTNCV-PQERELHKVNCFFSVFVAPVPEQYKILNKSC 114
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
83 YLLEVNIAITMCKKIAGDNENCILFQDDPKMKRMVFCIFVSSKPKFELKMKKKQC 138

RESULT 6
Q9BXP9 PRELIMINARY; PRT; 140 AA.
AC Q9BXP9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=21010502; PubMed=11144350;
RA Taupin P.H., Ray J., Fischer W.H., Suhr S.T., Hakansson K., Grubb A.,
RA Gage F.H.;
RT "PGF-2-Responsive neural stem cell proliferation requires Ccg, a novel
RT autocrine/paracrine cofactor";
RL Neuron 28:385-397(2000).
DR EMBL; AF311741; AAG40283.1; -.
DR HSSP; P01034; 1G96.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT CHAIN 21 140 CYSTATIN C.
FT VARIANT 16 16 A -> G.
FT VARIANT 84 84 L -> F.
SQ SEQUENCE 140 AA; 15517 MW; 3A563406D5D8D785 CRC64;

Query Match 26.0%; Score 166.5; DB 11; Length 140;
Best Local Similarity 33.3%; Pred. No. 6e-10;
Matches 35; Conservative 24; Mismatches 43; Indels 3; Gaps 2;

QY 15 EVMAVENYAKDSLQWITDQYNKESDDKYHFRIFVLKVRQVTDHLEHNLNVEQWTTCC 74
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
35 EADANEQGVRRALDFAVSEYKNGSNDYHSGRAIQVVRARKQLVAGVNYFLDVEGRTTCT 94

QY 75 KPET--TNC-VQERELHKVNCFFSVFVAPVPEQYKILNKSCSS 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
95 KSTQNLDCPFHQPHLMKALCSFQIYSVFWGTHSLTNFNSCN 139

RESULT 7
Q80V72 PRELIMINARY; PRT; 140 AA.
ID Q80V72
```

```
AC Q80V72;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cystatin-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenken C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schenken C.M., Schuler G.D.,
RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaracne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048646; AAH48646.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C/M.
DR Pfam; PF00031; cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 140 AA; 16199 MW; 32633E99C4697DA0 CRC64;

Query Match 23.4%; Score 149.5; DB 11; Length 140;
Best Local Similarity 30.4%; Pred. No. 3.8e-08;
Matches 31; Conservative 27; Mismatches 41; Indels 3; Gaps 2;

QY 15 EVMAVENYAKDSLQWITDQYNKESDDKYHFRIFVLKVRQVTDHLEHNLNVEQWTTCC 74
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
37 KAMSKKN-INSTLHFFIRSNASNDYLYVQVKLIQGMQLTGTGVEYLVTVKIGRTCK 95

QY 75 KPET--TNCVPQERELHKVNCFFSVFVAPVPEQYKILNKSC 114
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
96 KNETKASCPLQSSKLCKSLICKSLIYSVPMWNYQLMNNSC 137

RESULT 8
Q98SR4 PRELIMINARY; PRT; 112 AA.
ID Q98SR4;
AC Q98SR4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin (Fragment).
OS Acipenser sinensis (Chinese sturgeon).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=61970;
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 70.1071 Seconds  
(without alignments)  
526.560 Million cell updates/sec

Title: US-09-941-314-4  
Perfect score: 640  
Sequence: 1 LPYQARKKFLSVHEVWAVE.....VFAPVFEQYKILNKSCSSD 117

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:\*  
1: sp\_archea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvrius:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	640	100.0	138	4 Q8WXU6	Q8wxu6 homo sapien
2	411.5	64.3	103	4 Q8WXU5	Q8wxu5 homo sapien
3	360	56.2	139	11 Q8K5A3	Q8k5a3 rattus norv
4	188.5	29.5	141	11 Q9DAP1	Q9dap1 mus musculu
5	188.5	29.5	141	11 Q80ZN5	Q80zn5 mus musculu
6	166.5	26.0	140	11 Q9EPX9	Q9epx9 mus musculu
7	149.5	23.4	140	11 Q80Y72	Q80y72 mus musculu
8	139	21.7	112	13 Q98SR4	Q98sr4 acipenser s
9	139	21.7	112	13 Q98SR3	Q98sr3 acipenser s
10	122	19.1	130	11 Q8VIH8	Q8vih8 rattus norv
11	121	18.9	130	11 Q9CX46	Q9cx46 mus musculu
12	121	18.9	130	11 Q8VII3	Q8vii3 mus musculu
13	114.5	17.9	128	11 Q9DAN8	Q9dan8 mus musculu
14	113	17.7	148	11 Q9JW84	Q9jw84 mus musculu
15	112.5	17.6	146	11 Q8K397	Q8k397 mus musculu
16	112.5	17.6	149	11 Q8VHC1	Q8vhc1 rattus norv

17	112.5	17.6	149	11 Q9D1B1	Q9d1b1 mus musculu
18	109	17.0	167	4 Q7Z4J8	Q7z4j8 homo sapien
19	105	16.4	128	11 Q8VII2	Q8vii2 rattus norv
20	98.5	15.4	144	13 Q8JFUS	Q8jfu5 brachydanio
21	98	15.3	167	11 Q9QWLS	Q9qw15 mus musculu
22	95.5	14.9	109	5 Q9TY65	Q9ty65 onchocerca
23	92.5	14.5	161	5 O16159	O16159 brugia mala
24	90.5	14.1	148	5 Q9NH95	Q9nh95 litomosoides
25	86	13.4	133	11 Q9D264	Q9d264 mus musculu
26	84.5	13.2	125	5 Q25620	Q25620 onchocerca
27	84	13.1	127	5 P90698	P90698 brugia mala
28	84	13.1	425	3 Q12700	Q12700 debaryomyce
29	82.5	12.9	157	5 Q17108	Q17108 acanthochei
30	82	12.8	127	5 Q9U9A1	Q9u9a1 onchocerca
31	80.5	12.6	133	5 Q8MVB6	Q8mvb6 ixodes scap
32	80	12.5	498	5 O16454	O16454 caenorhabdi
33	79.5	12.4	2773	5 Q8IBF4	Q8ibf4 plasmodium
34	79	12.3	430	11 Q63581	Q63581 rattus norv
35	78	12.2	996	4 Q8NDM7	Q8ndm7 homo sapien
36	77.5	12.1	115	6 Q95K43	Q95k43 macaca fasc
37	76.5	12.0	462	13 Q7ZY91	Q7zy91 xenopus lae
38	75.5	11.8	209	5 Q815X0	Q815x0 plasmodium
39	75.5	11.8	465	13 Q801E5	Q801e5 xenopus lae
40	75	11.7	140	6 Q7YRP6	Q7yrp6 sus scrofa
41	75	11.7	400	4 Q8WXX6	Q8wxk6 homo sapien
42	75	11.7	400	4 Q8WXX7	Q8wxk7 homo sapien
43	75	11.7	423	11 P70517	P70517 rattus norv
44	75	11.7	891	5 Q8IB61	Q8ib61 plasmodium
45	74.5	11.6	658	16 Q81160	Q81160 bacillus ce

# ALIGNMENTS

RESULT 1	Q8WXU6	PRELIMINARY;	PRT;	138 AA.
ID	Q8WXU6	PRELIMINARY;	PRT;	138 AA.
AC	Q8WXU6			
DT	01-MAR-2002 (TREMBlrel. 20, Created)			
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)			
DT	01-JUN-2003 (TREMBlrel. 24, Last annotation update)			
DE	SC13.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;			
RT	"SC13: A novel epididymal specific member of the cystatin family.";			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF335480; AAL71991.1; -			
DR	GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR00010; Cystatin.			
DR	Pfam; PF00031; cystatin; 1.			
DR	SMART; SM00043; CY; 1.			
SQ	SEQUENCE 138 AA; 16506 MW; B49440ACA3585C64 CRC64;			

Query Match	100.0%;	Score 640;	DB 4;	Length 138;
Best Local Similarity	100.0%;	Pred. No. 2.9e-60;		
Matches 117;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1 LPYQARKKFLSVHEVWAVENYAKDSLOWITDQYNKESDDKYHFRIPRVLKQVRQVTDHL 60			
Db	22 LPYQARKKFLSVHEVWAVENYAKDSLOWITDQYNKESDDKYHFRIPRVLKQVRQVTDHL 81			
Qy	61 EYHLNVEMQWTTCKPQETTCVQRELHKQVNCFFSVPAWPFQYKILNKSCSSD 117			
Db	82 EYHLNVEMQWTTCKPQETTCVQRELHKQVNCFFSVPAWPFQYKILNKSCSSD 138			

RESULT 2  
Q8WXU5



Search completed: March 18, 2004, 14:16:08  
Job time : 13.9286 secs

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Rameay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [7]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Pahay J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [9]  
RP SEQUENCE OF 27-146.  
RX MEDLINE=82222268; PubMed=6283552;  
RA Grubb A., Loeffberg H.;  
RT "Human gamma-trace, a basic microprotein: amino acid sequence and  
presence in the adenohypophysis.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027(1982).  
RN [9]  
RP SEQUENCE OF 27-73.  
RX MEDLINE=84110059; PubMed=6662498;  
RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,  
RA Machleidt W.;  
RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
of cystatin from chicken egg white.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).  
RN [10]  
RP SEQUENCE OF 27-76.  
RX MEDLINE=84128015; PubMed=6365094;  
RA Brzin J., Popovic T., Turk V.;  
RT "Human cystatin, a new protein inhibitor of cysteine proteinases.";  
RL Biochem. Biophys. Res. Commun. 118:103-109(1984).  
RN [11]  
RP DISULFIDE BONDS.  
RA Grubb A., Loeffberg H., Barrett A.J.;  
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
cystatin.";  
RL FEBS Lett. 170:370-374(1984).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.  
RX MEDLINE=21171909; PubMed=11276250;  
RA Janowski R., Kozak M., Jankowska E., Grzonka Z., Grubb A.,  
RA Abrahamson M., Jaskolski M.;  
RT "Human cystatin C, an amyloidogenic protein, dimerizes through  
three-dimensional domain swapping.";  
RL Nat. Struct. Biol. 8:316-320(2001).  
RN [13]  
RP VARIANT GLN-94.  
RX MEDLINE=92316504; PubMed=1352269;

RA Abrahamson M., Jonsdottir S., Olafsson I., Jenson O., Grubb A.;  
RT "Hereditary cystatin C amyloid angiopathy: identification of the  
disease-causing mutation and specific diagnosis by polymerase chain  
reaction based analysis.";  
RL Hum. Genet. 89:377-380(1992).  
CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
CC thought to serve an important physiological role as a local  
CC regulator of this enzyme activity.  
CC -!- SUBUNIT: Homodimer.  
CC -!- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,  
CC vas deferens, brain, thymus, and ovary and the lowest in the  
CC submandibular gland.  
CC -!- DISEASE: Defects in CST3 are a cause of hereditary cerebral  
CC hemorrhage with amyloidosis (HCHWA) [MIM:105150]; also known as  
CC cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis  
CC Icelandic type. HCHWA is characterized by a thickening of the  
CC cerebral arteries walls with deposition of material with the  
CC characteristics of amyloid.  
CC -!- SIMILARITY: Belongs to the cystatin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X05607; CAA29096.1; -;  
CC EMBL; X52255; CAA36497.1; -;  
CC EMBL; M27891; AAA52164.1; -;  
CC EMBL; M27889; AAA52164.1; JOINED.  
CC EMBL; M27890; AAA52164.1; JOINED.  
CC EMBL; X61681; CAA43856.2; -;  
CC EMBL; X61682; CAA43856.2; JOINED.  
CC EMBL; X61683; CAA43856.2; JOINED.  
CC EMBL; AF319564; AAK11570.1; -;  
CC EMBL; AL121894; CAC05424.1; -;  
CC EMBL; BC013083; AAH13083.1; -;  
CC PIR; S10216; UDHU.  
CC PDB; 1G96; 06-APR-01  
CC Genew; HGNC:2475; CST3.  
CC MIM; 604312; -;  
CC MIM; 105150; -;  
CC InterPro; IPR000010; Cystatin.  
CC Pfam; PF00031; cystatin; 1.  
CC SMART; SM00043; Cy; 1.  
CC PROSITE; PS00287; CYSTATIN; 1.  
CC KW Thiol protease inhibitor; Amyloid; Signal; Disease mutation;  
CC Polymorphism; 3D-structure.  
CC FT SIGNAL 1 26  
CC FT CHAIN 27 146 CYSTATIN C.  
CC FT ACT\_SITE 37 37 REACTIVE SITE.  
CC FT SITE 81 85 SECONDARY AREA OF CONTACT.  
CC FT DISULFID 99 109  
CC FT DISULFID 123 143  

Query Match 24.2%; Score 155; DB 1; Length 146;  
Best Local Similarity 33.7%; Pred. No. 5.6e-09;  
Matches 34; Conservative 21; Mismatches 42; Indels 4; Gaps 3;  
QY 20 ENYAKDSLOWITDQYNKESDDKYHFRIVLKVQVQVTDHLEYHLNVEMQWTTCK--PE 77  
DB 46 EGVREALDFAVGEYKNKASNDMYHSALQVVRARKQIVAGVNYFLDELGRTTCTKTPN 105  
QY 78 TTNC-VQERELHKQVNCFFSFAVPWFQYKILNKSCSSD 117  
DB 106 LDCPPHDPHLKRKAFCSFQIVAVPW-OGTMTLSKSTCOD 145

RESULT 15  
CYTC\_MACMU  
ID -CYTC\_MACMU STANDARD; PRT; 146 AA.

RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahm D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Hunt A.R., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leivaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConnachie L.J., McIlroy K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,  
 RA Suze C.D., Smith M.L., Soderlund C., Steward C.A., Suleston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 CC -I- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -I- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC  
 DR EMBL: AL096677; CAC03440.2; -  
 DR HSP: P01038; J090  
 DR Genew; HGNC:15958; CST11.  
 DR InterPro; IPR000010; Cystatin.  
 DR PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
 KW Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 165 CYSTATIN-LIKE 1.  
 FT SITE 93 97 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 111 121 BY SIMILARITY.  
 FT DISULFID 134 154  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 165 AA; 19312 MW; 9D6D685875DAEEA CRC64;  
 Query Match 24.6%; Score 157.5; DB 1; Length 165;  
 Best Local Similarity 26.2%; Pred. No. 3.6e-09;  
 Matches 32; Conservative 30; Mismatches 37; Indels 23; Gaps 3;  
 QY 15 EVMAVENYAKDSQWITDQYKNSDDKYHFRIFRLVKVQ-----R 54  
 Db 34 KLSKKN-MNSTLNFFIQSTNNASNDYLRVRLRSQMRVSHWMLGVHNTSDSR 92  
 QY 55 QVTDHLEYHLNMQWTTQKPTTN--CVFQRELHKQVNCFFSFAVPWFQYKILNK 112  
 Db 93 QLITGVYEIVTVKIGWTCKKNDNTSSSCPLQSKLRKSLICESLIYTMPWVFLWN 152  
 QY 113 SC 114  
 Db 153 SC 154

RESULT 14  
 CYTC\_HUMAN STANDARD; PRT; 146 AA.  
 AC P01034;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)  
 DE (Post-gamma-globulin).  
 GN CST3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=87219149; PubMed=3495457;  
 RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;  
 RT "Molecular cloning and sequence analysis of cDNA coding for the  
 RT precursor of the human cysteine proteinase inhibitor cystatin C.";  
 RL FEBS Lett. 216:229-233(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=90303202; PubMed=2363674;  
 RA Abrahamson M., Olafsson I., Paledottir A., Ulvsbaeck M., Lundwall A.,  
 RA Jenson O., Grubb A.;  
 RT "Structure and expression of the human cystatin C gene.";  
 RL Biochem. J. 268:287-294(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A. (HCHWA VARIANT).  
 RC TISSUE=Brain;  
 RX MEDLINE=89235594; PubMed=2541223;  
 RA Levy E., Lopez-Otin C., Ghiso J., Gelnert D., Frangione B.;  
 RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is  
 RT related to a mutation in the cystatin C gene, an inhibitor of  
 RT cysteine proteases";  
 RL J. Exp. Med. 169:1771-1778(1989).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89350949; PubMed=2764935;  
 RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,  
 RA Isemura S., Sanada K.;  
 RT "The human cystatin C gene (CST3) is a member of the cystatin gene  
 RT family which is localized on chromosome 20.";  
 RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Dickinson D.P., Hewett-Emmett D., Thiesse M.;  
 RT "Acquisition of complex patterns of differential expression in  
 RT epithelial cell populations during the evolution of type 2 cystatin  
 RT genes.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahm D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leivaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConnachie L.J., McIlroy K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA



Gene 152:221-226 (1995).

[3]  
SEQUENCE FROM N.A.  
RC STRAIN=ILS, and ISS;  
RX MEDLINE=21363810; PubMed=11471062;  
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,  
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikeia J.M.;  
RA "High-throughput sequence identification of gene coding variants  
RT within alcohol-related QTLs";  
RT Mamm. Genome 12:657-663 (2001).  
RL [4]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).  
CC -I- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
CC thought to serve an important physiological role as a local  
CC regulator of this enzyme activity.  
CC -I- SIMILARITY: Belongs to the cystatin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL; M59470; AAA63298.1; --  
DR EMBL; UI0098; AAB41056.1; --  
DR EMBL; AF483486; AAL90760.1; --  
DR EMBL; AF483487; AAL90761.1; --  
DR EMBL; BC002072; AAH02072.1; --  
DR PIR; A36163; A36163.  
DR HSSP; P01034; IG96.  
DR MGD; MGI:102519; Cst3.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00031; cystatin; 1.  
DR SMART; SM00043; Cy; 1.  
DR PROSITE; PS00287; CYSTATIN; 1.  
DR Thiol protease inhibitor; Signal.  
KW FT SIGNAL 1 20  
FT CHAIN 21 140 CYSTATIN C.  
FT ACT SITE 31 31 REACTIVE SITE.  
FT SITE 75 79 SECONDARY AREA OF CONTACT.  
FT DISULFID 93 103 BY SIMILARITY.  
FT DISULFID 117 137 BY SIMILARITY.  
FT FT CONFLICT 16 16 A -> G (IN REF. 1).  
FT FT CONFLICT 84 84 L -> F (IN REF. 1).  
SQ SEQUENCE 140 AA; 15531 MW; 3A563406D58D0F5 CRC64;  
Query Match 25.1%; Score 160.5; DB 1; Length 140;  
Best Local Similarity 32.4%; Freq. No. 1.5e-09;  
Matches 34; Conservative 24; Mismatches 44; Indels 3; Gaps 2;

Query Match 25.1%; Score 160.5; DB 1; Length 140;  
Best Local Similarity 32.4%; Pred. No. 1.5e-09;  
Matches 34; Conservative 24; Mismatches 44; Indels 3; Gaps 2;

Qy	15	EWAVAVENYAKDSLOWITDYKNESDDKTHFRIFRVLVKQVRQUTDHLVHLVEMQWTTCT	7
Db	35	EDADNEEGVRRALDFAVSGEYNNKSGNDAYHSRAIQVVRARQKQLVAGVNYFLDVMGRITTC	94
Qy	75	KPET--TNC-VQERELHKQVNCFFSFVAPWPFEOYKILNKSCSS	116
Db	95	KSGNTLUDTCFFHDQPHLMRKALCSFQIYSVPWKGTHSLTKFSCKN	139
RESULT 11			
CYTT_HUMAN			
ID	CYTT_HUMAN	STANDARD;	PRT; 141 AA.
AC	P09228; Q9UCQ7;		
DT	01-MAR-1989 (Rel. 10, Created)		
DT	01-MAR-1989 (Rel. 10, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Cystatin SA precursor (Cystatin S5).		
GN	CS72.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID:9606;		
RN	[1]		
SEQUENCE FROM N.A.			
RP	MEDLINE=88185836; PubMed=3446578;		
RX	Saitoh E., Kim H.-S., Smithies O., Maeda N.;		
RT	"Human cysteine-proteinase inhibitors: nucleotide sequence analysis		
RT	of three members of the cystatin gene family.";		
RL	Gene 61:329-338 (1987).		
RN	[2]		
SEQUENCE FROM N.A.			
RP	MEDLINE=21638749; PubMed=11780052;		
RX	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,		
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,		
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,		
RA	Beasley O.P., Bird C.P., Blake S.E., Bridgeman A.M., Brown A.J.,		
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,		
RA	Chapman J.C., Clamp M., Clark L.N., Clark S.Y., Clee C.M.,		
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,		
RA	Coulson A., Coville G.J., Deaman R., Dhani P.D., Dunn M.,		
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,		
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,		
RA	Hammond S., Haxley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,		
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,		
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,		
RA	Lehvaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,		
RA	Marsh V.I., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,		
RA	Milne S.A., Mistry A., Moore M.J.F., Mullikin J.C., Nickerson T.,		
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,		
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,		
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,		
RA	Skuce C.D., Smith M.B., Soderlund C., Steward C.A., Sulston J.E.,		
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,		
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,		
RA	Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,		
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,		
RA	Rogers J.;		
RT	"The DNA sequence and comparative analysis of human chromosome 20.";		
RL	Nature 414:865-871 (2001).		
RN	[3]		
SEQUENCE OF 21-40.			
RC	TISSUE=Saliva;		
RX	MEDLINE=92138674; PubMed=1778989;		
RA	Isemura S., Saitoh E., Sanada K., Minakata K.;		
RT	"Identification of full-sized forms of salivary (S-type) cystatins		
RT	(cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of		
RT	cystatin S) in human whole saliva and determination of phosphorylation		
RT	sites of cystatin S.";		
RL	J. Biochem. 110:648-654 (1991).		
RN	[4]		
SEQUENCE OF 25-141.			
RX	MEDLINE=88139220; PubMed=3436950;		
RA	Isemura S., Saitoh E., Sanada K.;		



OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Buffalo;  
 RX MEDLINE=90092122; PubMed=2689174;  
 RA Cole T., Dickson P.W., Ennard F., Averill F., Risbridger G.,  
 RA Gauthier F., Schreiber G.;  
 RT "The cDNA structure and expression analysis of the genes for the  
 RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin  
 RT in rat brain.";  
 RL Eur. J. Biochem. 186:35-42(1989).  
 RN [2]  
 RP SEQUENCE OF 8-127.  
 RX MEDLINE=90380276; PubMed=2400577;  
 RA Ennard F., Ennard A., Faucher D., Capony J.-P., Derancourt J.,  
 RA Brillard M., Gauthier F.;  
 RT "Rat cystatin C: the complete amino acid sequence reveals a site for  
 RT N-glycosylation.";  
 RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).  
 RN [3]  
 RP SEQUENCE OF 8-49.  
 RX MEDLINE=89313020; PubMed=3044831;  
 RA Ennard A., Ennard F., Faucher D., Gauthier F.;  
 RT "Two rat homologues of human cystatin C.";  
 RL FEBS Lett. 236:475-478(1988).  
 RN [4]  
 RP SEQUENCE OF 8-20.  
 RC TISSUE=Sertoli cells;  
 RX MEDLINE=92225121; PubMed=1563513;  
 RA Ennard A., Ennard F., Guillou F., Gauthier F.;  
 RT "Production of the cysteine proteinase inhibitor cystatin C by rat  
 RT Sertoli cells.";  
 RL FEBS Lett. 300:131-135(1992).  
 CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 CC thought to serve an important physiological role as a local  
 CC regulator of this enzyme activity. Known to inhibits cathepsin B,  
 CC H, and L.  
 CC -!- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X16957; CAA34831.1; -;  
 DR PIR; S07085; S07085.  
 DR PIR; S10587; S10587.  
 DR HSSP; P01034; I096.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; Cx; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 KW Thiol protease inhibitor; Signal.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 7  
 FT CHAIN 8 127 CYSTATIN C.  
 FT ACT\_SITE 18 18 REACTIVE SITE.  
 FT SITE 62 66 SECONDARY AREA OF CONTACT.  
 FT DISULFID 80 90 BY SIMILARITY.  
 FT DISULFID 104 124 BY SIMILARITY.  
 FT CONFLICT 25 25 A -> E (IN REF. 2).  
 SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925953 CRC64;  
 Query Match 26.8%; Score 171.5; DB 1; Length 127;  
 Best Local Similarity 33.3%; Pred. No. 9.9e-11;  
 Matches 35; Conservative 24; Mismatches 43; Indels 3; Gaps 2;  
 OX 15 EVMAVNYAKDSQWITDQYKESDDKHPRIFRLVKVRQVTDHLEYHLNVMQWTTQC 74  
 DB 22 EADASEGVQRALDFAVSEYKNGSDAYH5RAIQVVRARQLVAGINYILDVEMGTTCT 81

QY 75 KPET--TNC--VPQERELHKQVNCFFSVFAVPWFQYKILNKSCSS 116  
 DB 82 KSQTNLTNCFFHDQPHLMKALCSQIYSVPKGTHTLTKSCKN 126  
 RESULT 8  
 CVT\_CHICK  
 ID\_CVT\_CHICK STANDARD; PRT; 139 AA.  
 AC P01038;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin precursor (Egg-white cystatin).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90008873; PubMed=2793849;  
 RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;  
 RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,  
 RT and tissue distribution.";  
 RL J. Biol. Chem. 264:17164-17169(1989).  
 RN [2]  
 RP SEQUENCE OF 24-139.  
 RX MEDLINE=84178305; PubMed=6712597;  
 RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;  
 RT "Cystatin. Amino acid sequence and possible secondary structure.";  
 RL Biochem. J. 217:813-817(1984).  
 RN [3]  
 RP SEQUENCE OF 24-139.  
 RX MEDLINE=84110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,  
 RA Machleidt W.;  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 RT of cystatin from chicken egg white.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).  
 RN [4]  
 RP CHARACTERIZATION OF PROTEIN.  
 RX MEDLINE=83256421; PubMed=6409085;  
 RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Savers C.A.,  
 RA Santer D.C., Barrett A.J.;  
 RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved  
 RT purification from egg white, characterization, and detection in  
 RT chicken serum.";  
 RL Biochem. J. 211:129-138(1983).  
 RN [5]  
 RP DISULFIDE BONDS.  
 RA Grubb A., Loeffberg H., Barrett A.J.;  
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
 RT cystatin.";  
 RL FEBS Lett. 170:370-374(1984).  
 RN [6]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=89252033; PubMed=2721673;  
 RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,  
 RA Bode W.;  
 RT "The cysteine proteinase inhibitor chicken cystatin is a  
 RT phosphoprotein.";  
 RL FEBS Lett. 248:162-168(1989).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=89052676; PubMed=3191914;  
 RA Bode W., Eggh R., Musil D., Thiele U., Huber R., Karshikov A.,  
 RA Brzin J., Kos J., Turk V.;  
 RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and  
 RT its possible mode of interaction with cysteine proteinases.";  
 RL EMBO J. 7:2593-2599(1988).  
 RN [8]  
 RP STRUCTURE BY NMR.



[illegible]

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL; AL096677; CAC13170.1; -;  
 DR EMBL; AL096677; CAC17423.1; -;  
 DR HSSP; P01038; 1A90.  
 DR Genew; HGNC:15959; CST11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; Cy; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
 KW Thiol protease inhibitor; Signal; Alternative splicing.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 137  
 FT SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 93 101 BY SIMILARITY.  
 FT DISULFID 114 134 BY SIMILARITY.  
 FT CARBOHYD 131 131 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT VARSPLIC 76 110 Missing (in isoform 2).  
 FT FTID=VSP 001260.  
 SQ SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 640; DB 1; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-58;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKQVQVTDHLL 60  
 DB 21 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKQVQVTDHLL 80  
 QY 61 EYHLNEMQWTTCKPQETTCVQPERLHKQVNCFFSVFVAPVFPFQYKILNKSCSSD 117  
 DB 81 EYHLNEMQWTTCKPQETTCVQPERLHKQVNCFFSVFVAPVFPFQYKILNKSCSSD 137

## RESULT 2

CST11\_MOUSE STANDARD; PRT; 139 AA.  
 AC Q9D269;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin 11 precursor.  
 GN CST11.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadori K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleishmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,  
 RA Schral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi R.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."

RL Nature 409:685-690(2001).  
 CC -I- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -I- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL; AK020300; BAB32061.1; -;  
 DR HSSP; P01034; 1G96.  
 DR MGD; MGI:1925490; Cst11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; Cy; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
 KW Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT CHAIN 29 139  
 FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 94 102 BY SIMILARITY.  
 FT DISULFID 115 135 BY SIMILARITY.  
 FT CARBOHYD 134 134 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 139 AA; 16217 MW; P228D9815FA32640 CRC64;

Query Match 58.1%; Score 372; DB 1; Length 139;  
 Best Local Similarity 54.9%; Pred. No. 3.7e-31;  
 Matches 62; Conservative 29; Mismatches 22; Indels 0; Gaps 0;

QY 3 YQARKKTFSLVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKQVQVTDHLEY 62  
 DB 24 YQVQKRTFRIEVSALSSVKTLEYVTDYDNKESDLYNFRILKIMKQVTHLEY 83  
 QY 63 HLNEMQWTTCKPQETTCVQPERLHKQVNCFFSVFVAPVFPFQYKILNKSCS 115  
 DB 84 HITVEMQRTTCLKTETSLCDIQKGLHKKIQCVFSVYVAPVFPVFKILKQKNT 136

## RESULT 3

CST8\_MOUSE STANDARD; PRT; 142 AA.  
 AC P32766; O89102;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-  
 DE related epididymal specific protein) (Cystatin 8).  
 GN CST8 OR CR8.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H, and CD-1;  
 RX MEDLINE=99247899; PubMed=10229662;  
 RA Cornwell G.A., Hsia N., Sutton H.G.;  
 RT "Structure, alternative splicing and chromosomal localization of the  
 RT cystatin-related epididymal spermatogenic gene.";  
 RL Biochem. J. 340:85-93(1999).  
 RN [2]  
 RP SEQUENCE OF 4-142 FROM N.A.  
 RC TISSUE=Epididymis;  
 RX MEDLINE=93078799; PubMed=1280328;  
 RA Cornwell G.A., Orgebin-Crist M.-C., Hann S.R.;  
 RT "The CR8 gene: a unique testis-regulated gene related to the cystatin  
 RT family is highly restricted in its expression to the proximal region  
 RT of the mouse epididymis.";  
 RL Mol. Endocrinol. 6:1653-1664(1992).  
 CC -I- FUNCTION: Performs a specialized role during sperm development and

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 13.9286 Seconds  
(without alignments)  
437.389 Million cell updates/sec

Title: US-09-941-314-4  
Perfect score: 640  
Sequence: 1 LPYQARKKTFSLVHEVMAVE.....VFAPWFEQYKILNKSCSSD 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	640	100.0	137	CS11 HUMAN	Q9h112 homo sapien
2	372	58.1	139	CS11 MOUSE	P01042 mus musculus
3	214	33.4	142	CST8 MOUSE	P32766 mus musculus
4	199	31.1	142	CST8 RAT	O88959 rattus norv
5	188.5	29.5	142	CST8 HUMAN	O80676 homo sapien
6	174.5	27.3	148	CYT8 RHABIT	O97862 oryctolagus
7	171.5	26.8	127	CYT8 RAT	P14841 rattus norv
8	170.5	26.6	139	CYT CHICK	P01038 gallus gall
9	169.5	26.5	116	CYT COTUA	P81061 coturnix co
10	160.5	25.1	140	CYT8 MOUSE	P21450 mus musculus
11	160.5	25.1	141	CYT8 HUMAN	P09228 homo sapien
12	160	25.0	146	CYT8 SAISC	O19093 saimiri sci
13	157.5	24.6	165	CSTL HUMAN	Q9h114 homo sapien
14	155	24.2	146	CYT8 HUMAN	P01034 homo sapien
15	155	24.2	146	CYT8 MACMU	O19092 macaca mula
16	153.5	24.0	148	CYT8 BOVIN	P01035 bos taurus
17	146.5	22.9	141	CYT8 HUMAN	P01036 homo sapien
18	143	22.3	129	CYT CYPCA	P35481 cyprinus ca
19	139.5	21.8	141	CYT8 HUMAN	P01037 homo sapien
20	135	21.1	147	CST9 HUMAN	O9h4c1 homo sapien
21	130.5	20.4	142	CYT8 HUMAN	P28325 homo sapien
22	129.5	20.2	141	CYT8 RAT	P19313 rattus norv
23	127	19.8	111	CYT BITAR	P08935 bitis ariet
24	127	19.8	130	CYT ONCMY	Q91195 oncorhynch
25	127	19.8	137	CST9 MOUSE	Q920h6 mus musculus
26	126	19.7	130	CYT ONCKE	O98967 oncorhynch
27	121.5	19.0	149	CYT8 HUMAN	O15828 homo sapien
28	109	17.0	145	CYT8 HUMAN	O76096 homo sapien
29	98.5	15.4	434	KNL2 BOVIN	P01047 bos taurus
30	98.5	15.4	619	KNH2 BOVIN	P01045 bos taurus
31	98	15.3	144	CYT8 BOVIN	O89098 mus musculus
32	96.5	15.1	436	KNL1 BOVIN	P01046 bos taurus
33	96.5	15.1	621	KNH1 BOVIN	P01044 bos taurus

RESULT 1				
CS11_HUMAN	STANDARD;	PRT;	137	AA.
AC	Q9H112; Q9H113;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DE	Cystatin 11 precursor.			
GN	CST11 OR CST8L.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RX	MEDLINE=21638749; PubMed=11780052;			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,			
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,			
RA	Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,			
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,			
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,			
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,			
RA	Lehvaeslao M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,			
RA	Marsh V.L., Martin S.L., McConachie L.J., McMay K., McMurray A.A.,			
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,			
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Beck A.I.,			
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,			
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,			
RA	Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,			
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,			
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,			
RA	Whithead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,			
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,			
RA	Rogers J.;			
RT	"The DNA sequence and comparative analysis of human chromosome 20.;"			
RL	Nature 414:865-871(2001).			
CC	-1- SUBCELLULAR LOCATION: Secreted (Potential).			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=1;			
CC	isoId=Q9H112-1; Sequence=Displayed;			
CC	Name=2;			
CC	isoId=Q9H112-2; Sequence=VSP_001260;			
CC	Note=No experimental confirmation available;			
CC	-1- SIMILARITY: Belongs to the cystatin family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			

P22085 onchocerca  
P01042 homo sapien  
P08934 rattus norv  
Q87rl6 vibrrio para  
P08932 rattus norv  
P01048 rattus norv  
O08677 mus musculus  
O74232 kluyveromyc  
P44689 haemophilus  
Q7mmw8 vibrrio vuln  
Q06430 homo sapien  
P29717 candida alb

ALIGNMENTS



Search completed: March 18, 2004, 14:23:12  
Job time : 22.5179 secs



A;Cross-references: GB:J03870; NID:G337751; PIDN:AAA60299.1; PID:G337752  
R;Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.  
Biochem. Hoppe-Seyler 369, 191-197, 1988  
A;Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily  
A;Reference number: S02489; MUID:89076505; PMID:3202964  
A;Accession: S02489  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 21-141 <SA2>  
R;Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.  
Gene 61, 329-338, 1987  
A;Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three mem  
A;Reference number: A91589; MUID:88185836; PMID:3448578  
A;Accession: A29632  
A;Molecule type: DNA  
A;Residues: 1-86, 'I', 88-141 <SA1>  
R;Isemura, S.; Saitoh, E.; Sanada, K.  
FEBS Lett. 196, 145-149, 1986  
A;Title: Characterization of a new cysteine proteinase inhibitor of human saliva, cystat  
A;Reference number: A01273; MUID:86164938; PMID:3514272  
A;Accession: A01273  
A;Molecule type: protein  
A;Residues: 29-141 <SE>  
R;Ramasubbu, N.; Reddy, M.S.; Bergey, E.J.; Haraszthy, G.G.; Soni, S.D.; Levine, M.J.  
Biochem. J. 280, 341-352, 1991  
A;Title: Large-scale purification and characterization of the major phosphoproteins and  
A;Reference number: S19279; MUID:92082469; PMID:1747107  
A;Accession: S19279  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 21-55 <RAM>  
C;Comment: Human saliva appears to contain several cysteine proteinase inhibitors that a  
ences. Cystatin SN, with a pI of 7.5, is a much better inhibitor of papain and dipeptidyl  
C;Genetics:  
A;Gene: GDB:CST1  
A;Cross-references: GDB:119815; OMIM:123855  
A;Map position: 20p11.2-20p11.2  
C;Superfamily: cystatin; cystatin homology  
C;Keywords: cysteine proteinase inhibitor; extracellular protein; saliva  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-141/Product: cystatin SA-I #status experimental <MAT1>  
F;29-141/Product: cystatin SN #status experimental <MAT2>  
F;30-141/Domain: cystatin homology <CVS>  
F;76-80/Region: inhibitory #status predicted  
F;94-104,118-138/Disulfide bonds: #status predicted

Query Match 21.8%; Score 139.5; DB 1; Length 141;  
Best Local Similarity 29.6%; Pred. No. 8.9e-07;  
Matches 29; Conservative 21; Mismatches 45; Indels 3; Gaps 2;  
QY 20 ENYAKDSLWITDQYNKESDKYHFRIFRVLKVRQVTDHLEYHLNVEMQWTTCCQK--PE 77  
Db 41 DEWQVRAHFAISYNKATKDDYRRPRLVLRARQQTGGVNYFFDEVEGRITCTKSQPN 100  
QY 78 TTNCVQPER-ELHKQVRCFFSVFAPVWPEQYKILNKSC 114  
Db 101 LDTCAFHEQPELQKQLCSFEIYEVWPENRRSLVKSRC 138

RESULT 11  
JC2040  
cystatin - chum salmon  
N;Alternate names: cysteine proteinase inhibitor  
C;Species: Oncorhynchus keta (chum salmon)  
C;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 16-Jul-1999  
C;Accession: JC2040  
R;Koide, Y.; Noso, T.  
BioSci. Biotechnol. Biochem. 58, 164-169, 1994  
A;Title: The complete amino acid sequence of pituitary cystatin from chum salmon.  
A;Reference number: JC2040; MUID:94162738; PMID:7764512  
A;Accession: JC2040  
A;Molecule type: protein  
A;Residues: 1-111 <KOI>

A;Accession: A01270  
A;Molecule type: protein  
A;Residues: 27-131,'S',133-146 <GRU>  
R;Chiao, J.; Jensen, O.; Frangione, B.  
Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986  
A;Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland  
A;Reference number: A25434; MUID:86206076; PMID:3517880  
A;Accession: A25434  
A;Molecule type: protein  
A;Residues: 37-93,'O',95-146 <GHI>  
R;Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983  
A;Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin  
A;Reference number: S01461; MUID:84110059; PMID:6662498  
A;Accession: S12288  
A;Molecule type: protein  
A;Residues: 27-73 <TUR>  
R;Brzin, J.; Popovic, T.; Turk, V.  
Biochem. Biophys. Res. Commun. 118, 103-109, 1984  
A;Title: Human cystatin, a new protein inhibitor of cysteine proteinases.  
A;Reference number: A32732; MUID:84128015; PMID:6365094  
A;Accession: A32732  
A;Molecule type: protein  
A;Residues: 27-76 <BRZ>  
R;Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.  
Scand. J. Clin. Lab. Invest. 50, 85-93, 1990  
A;Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst  
A;Reference number: A60552; MUID:90193615; PMID:2315647  
A;Accession: A60552  
A;Molecule type: protein  
A;Residues: 27-49,'XX',52-64 <OLA>  
A;Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d  
e defective gene is not present in CSF but is found instead in amyloid deposits  
R;Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.  
Biol. Chem. Hoppe-Seyler 371, 575-580, 1990  
A;Title: Different forms of human cystatin C.  
A;Reference number: S10607; MUID:91025625; PMID:2222856  
A;Accession: S10607  
A;Molecule type: protein  
A;Residues: 27-53 <POP>  
A;Experimental source: urine, kidney disease  
A;Note: truncated forms with amino ends at positions 35 and 36 of the precursor were als  
R;Grubb, A.; Lofberg, H.; Barrett, A.J.  
FEBS Lett. 170, 370-374, 1984  
A;Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
A;Reference number: S01462  
A;Contents: annotation; disulfide bonds  
R;Berti, P.J.; Storer, A.C.  
Biochem. J. 302, 411-416, 1994  
A;Title: Local pH-dependent conformational changes leading to proteolytic susceptibility  
A;Reference number: S55305; MUID:94379969; PMID:8092991  
A;Accession: S55305  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 27-49;106-146 <BER>  
C;Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl  
f patients with certain autoimmune diseases.  
C;Comment: This protein is an inhibitor of cysteine proteinases and may serve an importa  
C;Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorr  
C;Genetics:  
A;Gene: GDB:CST3  
A;Cross-references: GDB:1119817; OMIM:105150  
A;Map position: 20p11.2-20p11.2  
A;Introns: 81/3, 119/3  
C;Superfamily: Cystatin; cystatin homology  
C;Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyprolin  
F;1-26/Domain: signal sequence #status predicted <SIG>  
F;27-146/Product: cystatin C #status experimental <MAT>  
F;35-146/Domain: cystatin homology <CVS>  
F;81-85/Region: inhibitory #status predicted  
F;23/Modified site: hydroxyproline (Pro) (partial) #status experimental  
F;99-109,123-143/Disulfide bonds: #status experimental

Query Match 24.2%; Score 155; DB 1; Length 146;  
Best Local Similarity 33.7%; Pred. NO. 2.7e-08;  
Matches 34; Conservative 21; Mismatches 42; Indels 4; Gaps 3;  
QY 20 ENYAKDSLQWITDQYNKESDDKYHFRIFVLKVKQROVTDHLEHYHLNVEQMWTTCQK--PE 77  
DB 46 ESGVRALDFAVGEYKNSNDMTHSALQVVRARKQIVAGVNYFLDVLGRITCTKTOPN 105  
QY 78 TTNC-VPQREHLHKQVNCFFSVFAVFPWFQYKILNKSCSSD 117  
DB 106 LNCPCPHDQPHLKRKAFCSQIYAVPW-QGTWTLSTKSTCOD 145  
RESULT 8  
UDU01  
cystatin - bovine  
N;Alternate names: thiol proteinase inhibitor  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 28-Feb-1986 #sequence\_revision 28-Feb-1986 #text\_change 06-Dec-1996  
C;Accession: A01271  
R;Hirado, M.; Tsunagaawa, S.; Sakiyama, F.; Niinobe, M.; Fujii, S.  
FEBS Lett. 186, 41-45, 1985  
A;Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase in  
A;Reference number: A01271; MUID:85231205; PMID:3891407  
A;Accession: A01271  
A;Molecule type: protein  
A;Residues: 1-112 <HIR>  
C;Superfamily: cystatin; cystatin homology  
C;Keywords: colostrum; cysteine proteinase inhibitor  
F;2-112/Domain: cystatin homology <CVS>  
F;48-52/Region: inhibitory #status predicted  
F;66-76,90-110/Disulfide bonds: #status predicted  
Query Match 24.0%; Score 153.5; DB 1; Length 112;  
Best Local Similarity 30.6%; Pred. NO. 2.8e-08;  
Matches 30; Conservative 23; Mismatches 42; Indels 3; Gaps 2;  
QY 20 ENYAKDSLQWITDQYNKESDDKYHFRIFVLKVKQROVTDHLEHYHLNVEQMWTTCQKPE 79  
DB 13 EEGVQALGFVAVSEFNKRSNDAYQSRVVRVVRARKQVSGMNYFLDVLGRITCTKSOAN 72  
QY 80 --NC-VPQREHLHKQVNCFFSVFAVFPWFQYKILNKSC 114  
DB 73 LDSCPFHPHQLKREKLCFSQVYVVPVPMNTINLVKFS 110  
RESULT 9  
UDU01  
cystatin S precursor - human  
N;Alternate names: cystatin SA-III; salivary acidic protein-1  
C;Species: Homo sapiens (man)  
C;Date: 25-Feb-1985 #sequence\_revision 08-Feb-1996 #text\_change 16-Jul-1999  
C;Accession: S17667; S16500; A01272; A29603; S19280; A56508  
R;Bobek, L.A.; Aguirre, A.; Levine, M.J.  
Biochem. J. 278, 627-635, 1991  
A;Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ and  
A;Reference number: S17667; MUID:91378918; PMID:1898352  
A;Accession: S17667  
A;Molecule type: mRNA  
A;Residues: 1-141 <BOB>  
A;Cross-references: EMBL:X54667; NID:G30365; PIDN:CAA38478.1; PID:G30366  
R;Lamkin, M.S.; Jensen, J.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.  
Arch. Biochem. Biophys. 288, 664-670, 1991  
A;Title: Salivary cystatin SA-III, a potential precursor of the acquired enamel pellicle,  
A;Reference number: S16500; MUID:91378515; PMID:1898055  
A;Accession: S16500  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 21-134,'D',136-141 <IHU>  
R;Isemura, S.; Saitoh, E.; Sanada, K.  
J. Biochem. 96, 489-498, 1984  
A;Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole sal  
A;Reference number: A91985; MUID:85054716; PMID:6501254

Db 80 LTNCPPHDPHLMRKALCSFQIYSPVWKGTHTLTKSSCKN 119

RESULT 5

B29632

cystatin SA precursor - human

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1989 #sequence revision 10-Jun-1989 #text\_change 16-Jul-1999

C:Accession: B29632, S02490, A41422, B27015

R:Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.

Gene 61, 329-338, 1987

A:Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three members

A:Reference number: A91589; MUID:88185836; PMID:3446578

A:Accession: B29632

A:Molecule type: DNA

A:Residues: 1-141 <SAI>

A:Cross-references: GB:M19673; GB:M19170; NID:G186403; PIDN:AAA36116.1; PID:G386826

A>Note: the authors translated the codon GAC for residue 129 as Asn

R:Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.

Biol. Chem. Hoppe-Seyler 369, 191-197, 1988

A:Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily

A:Reference number: S02489; MUID:89076505; PMID:3202964

A:Accession: S02490

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 21-141 <SA2>

R:Isemura, S.; Saitoh, E.; Sanada, K.

J. Biochem. 102, 693-704, 1987

A:Title: Characterization and amino acid sequence of a new acidic cysteine proteinase in

A:Reference number: A41422; MUID:88139220; PMID:3436950

A:Accession: A41422

A:Molecule type: protein

A:Residues: 25-141 <IS2>

R:Isemura, S.; Saitoh, E.; Sanada, K.; Ito, S.

In Cysteine Proteinases and Their Inhibitors, Turk, V., ed., pp.497-505, Walter de Gruyter

A:Title: Cystatin S and the related cysteine proteinase inhibitors in human saliva.

A:Reference number: A27015

A:Accession: B27015

A:Molecule type: protein

A:Residues: 25-134, 'D', '136-141 <IS2>

C:Genetics:

A:Gene: GDB:CST2

A:Cross-references: GDB:119816; OMIM:123856

A:Map position: 20p11.2-20p11.2

C:Superfamily: cystatin; cystatin homology

F:30-141/Domain: cystatin homology <CYS>

Query Match 25.1%; Score 160.5; DB 2; Length 141;

Best Local Similarity 31.5%; Pred. No. 7.3e-09;

Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDQNKSDDKYHFRIFVLKQVQVTDHLEYHLNVEMQWTTCC--PETTNCVP 83

Db 47 ALHFVISEYNKATEDEYRLLRLVLRAREQIVGVNVPFFDIEVGRITCTKSPNLDTCAP 106

QY 84 QER-ELHKQVNCFFSVFAVPWFQYKILNKSC 114

Db 107 HQPELOKQKLCSCFIYVFPWEDRMSLVNSRC 138

RESULT 6

A36163

cystatin C precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 16-Jul-1999

C:Accession: A36163

R:Solem, M.; Rawson, C.; Lindburg, K.; Barnes, D.

Biochem. Biophys. Res. Commun. 172, 945-951, 1990

A:Title: Transforming growth factor beta regulates cystatin C in serum-free mouse embryo

A:Reference number: A36163; MUID:91054522; PMID:2241983

A:Accession: A36163

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140 <SOL>

A:Cross-references: EMBL:M59470; NID:G192911; PIDN:AAA63298.1; PID:G192912

C:Superfamily: cystatin; cystatin homology

F:29-140/Domain: cystatin homology <CYS>

F:93-103,117-137/Dialfide bonds: #status predicted

Query Match 24.5%; Score 156.5; DB 2; Length 140;

Best Local Similarity 31.4%; Pred. No. 1.8e-08;

Matches 33; Conservative 24; Mismatches 45; Indels 3; Gaps 2;

QY 15 EYMAVENYAKDSLOWITDQNKSDDKYHFRIFVLKQVQVTDHLEYHLNVEMQWTTCC 74

Db 35 EADANEVGVRRALDFAVSEYKNGSDAYHSRAIQVVRARQQLVAGVNYFFDVEGMGRITCT 94

QY 75 KPET--TNC-VPQERELHKQVNCFFSVFAVPWFQYKILNKSCSS 116

Db 95 KSQTNLTDCPPHDPHLMRKALCSFQIYSPVWKGTHSLTKFSCKN 139

RESULT 7

UDRU

cystatin C precursor [validated] - human

N:Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma I

C:Species: Homo sapiens (man)

C>Date: 06-Jul-1982 #sequence revision 31-Mar-1991 #text change 08-Dec-2000

C:Accession: S10216; S00004; JLO095; A33400; S02751; A01270; A25434; S12288; A32732; A60

R:Abrahamson, M.; Olafsson, I.; Palsdottir, A.; Ulvbaeck, M.; Lundwall, A.; Jensen, O.

Biochem. J. 268, 287-294, 1990

A:Title: Structure and expression of the human cystatin C gene.

A:Reference number: S10216; MUID:90303202; PMID:2363674

A:Accession: S10216

A:Molecule type: DNA

A:Residues: 1-146 <AB1>

R:Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.

FEBS Lett. 216, 229-233, 1987

A:Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the

A:Reference number: S00004; MUID:87219149; PMID:3495457

A:Accession: S00004

A:Molecule type: mRNA

A:Residues: 1-146 <AB2>

A:Cross-references: EMBL:X05607; NID:G30371; PIDN:CAA29096.1; PID:G755738

R:Levy, E.; Lopez-Otin, C.; Ghiso, J.; Geltner, D.; Frangione, B.

J. Exp. Med. 169, 1771-1778, 1989

A:Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a

A:Reference number: JLO095; MUID:89235594; PMID:2541223

A:Accession: JLO095

A:Molecule type: DNA

A:Residues: 1-146 <LEV>

A:Cross-references: GB:X61681; NID:G30367; PIDN:CAA43856.2; PID:G4490944

A>Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit

e)

R:Saitoh, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada, K.

Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989

A:Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which i

A:Reference number: A33400; MUID:89350949; PMID:2764935

A:Accession: A33400

A:Molecule type: DNA

A:Residues: 1-24, 'T', '25-146 <SAI>

A:Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:G181385; PIDN:AAA52164.1; PID:G1

R:Ghiso, J.; Cowan, N.; Frangione, B.

Biol. Chem. Hoppe-Seyler 369, 205-208, 1988

A:Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron

A:Reference number: S02751; MUID:89076507; PMID:3264504

A:Accession: S02751

A:Molecule type: DNA

A:Residues: 82-119 <GH2>

A:Cross-references: EMBL:M27769

A>Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown i

R:Grubb, A.; Lofberg, H.

Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982

A:Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in the

A:Reference number: A01270; MUID:82222268; PMID:6283552

A;Residues: 24-139 <AUE>  
R;Luber, B.; Krieglstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.  
FEBS Lett. 248, 162-168, 1989  
A;Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.  
A;Reference number: S04008; MUID:89252033; PMID:2721673  
A;Accession: S04008  
A;Molecule type: protein  
A;Residues: 97-114 <UAB>  
R;Colella, R.; Bird, J.W.C.  
Gene 130, 175-181, 1993  
A;Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping t  
A;Reference number: JN0789; MUID:93366172; PMID:8359684  
A;Accession: JN0789  
A;Molecule type: DNA  
A;Residues: 1-139 <CO2>  
A;Cross-references: GB:M95725  
A;Note: authors failed to translate the codon for residue 115-Tyr  
C;Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases ;  
C;Genetics:  
A;Gene: Csn  
A;Introns: 76/3; 114/3  
C;Superfamily: cystatin; cystatin homology  
C;Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-139/Product: cystatin, long form #status experimental <CYLP>  
F;30-139/Domain: cystatin homology <CYS>  
F;32-139/Product: cystatin, short form #status experimental <CYSF>  
F;76-80/Region: inhibitory #status predicted  
F;94-104,118-138/Disulfide bonds: #status experimental  
F;103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 26.6%; Score 170.5; DB 1; Length 139;  
Best Local Similarity 37.0%; Pred. No. 7.3e-10;  
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWIDQYNKESDDKYHFRFLRVLKVRQVTDHLEHLNVEMQWTTCKP--ETTNC-V 82  
DB 47 ALGFAMAEYNRASNDKYSRRVVRVISAKRLQVSGIKYILQVEIGRTTCKPSGDLQSCF 106  
QY 83 PQRELHKQVNCFFSPVAPVWFQYKIILKSC 114  
DB 107 HDEPEAKYTTCTFFVVISIPLWNLKLESKC 138

RESULT 4  
S10587  
Cystatin C - rat  
C;Species: Rattus sp. (rat)  
C;Date: 21-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 16-Jul-1999  
C;Accession: S10587  
R;Esnard, F.; Esnard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Brillard, M.; Gauthi  
Biol. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990  
A;Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylat  
A;Reference number: S10587; MUID:90380276; PMID:2400577  
A;Accession: S10587  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-120 <ESN>  
A;Note: 43-Asn was also found  
A;Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala  
C;Superfamily: cystatin; cystatin homology  
F;9-120/Domain: cystatin homology <CYS>

Query Match 26.2%; Score 167.5; DB 2; Length 120;  
Best Local Similarity 33.0%; Pred. No. 1.2e-09;  
Matches 33; Conservative 24; Mismatches 40; Indels 3; Gaps 2;

QY 20 ENYAKDSLOWITDQYNKESDDKYHFRFLRVLKVRQVTDHLEHLNVEMQWTTCKPET- 78  
DB 20 EEGVQRALDPAVSEYNGSDNAYHSRALQVVRKQLVAGINYYLDVEMGRITCTKQTIN 79  
QY 79 -TNC-VQERELHKQVNCFFSPVAPVWFQYKIILKNCSS 116

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 22.5179 Seconds  
(without alignments)  
499.799 Million cell updates/sec

Title: US-09-941-314-4  
Perfect score: 640  
Sequence: 1 LPQARKKTFLSVHEVWAVE.....VFAVPWFQYKILNKSCSSD 117

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues 283366  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: Pir1:\*  
2: Pir2:\*  
3: Pir3:\*  
4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	214	33.4	139	2 A45361	cystatin-related e
2	171.5	26.8	127	2 S07085	cystatin C precurs
3	170.5	26.6	139	1 UDCH	cystatin precursor
4	167.5	26.2	120	2 S10587	cystatin C - rat
5	160.5	25.1	141	2 B29632	cystatin SA precu
6	156.5	24.5	140	2 A36163	cystatin C precurs
7	155	24.2	146	1 UDHU	cystatin C precurs
8	153.5	24.0	112	1 UDBO	cystatin - bovine
9	146.5	22.9	141	1 UDHUP1	cystatin S precurs
10	139.5	21.8	141	1 UDHUP2	cystatin SN precu
11	134	20.9	111	1 JC2040	cystatin - chum sa
12	131.5	20.5	142	2 A47142	cystatin D precurs
13	129.5	20.2	141	2 JQ1470	cystatin S precurs
14	127	19.8	111	2 A28793	cystatin - puff ad
15	126	19.7	132	2 JC4918	cystatin precursor
16	102.5	16.0	133	2 JC4536	cystatin precursor
17	98.5	15.4	434	1 KGBOL2	kininogen, LMW II
18	98.5	15.4	619	1 KGBOH2	kininogen, HMW II
19	96.5	15.1	436	1 KGBOL1	kininogen, LMW I p
20	96.5	15.1	621	1 KGBOH1	kininogen, HMW I p
21	95.5	14.9	162	2 A43428	onchocystatin - ne
22	93.5	14.6	427	1 KGHUL1	kininogen, LMW pre
23	93.5	14.6	644	1 KGHUL1	kininogen, HMW pre
24	87	13.6	433	2 A28055	K-kininogen, LMW I
25	87	13.6	639	2 A28486	K-kininogen, HMW I p
26	85.5	13.4	430	2 A23897	major acute phase
27	81	12.7	430	2 B28055	T-kininogen, LMW I
28	80	12.5	498	2 T31871	hypothetical prote
29	79	12.3	430	1 KGRTT1	T-kininogen I prec

30	77	12.0	430	2 D64151	hypothetical prote
31	75	11.7	400	2 A46297	beta-1,6-N-acetylgl
32	75	11.7	423	1 KGRTM	major acute phase
33	74	11.6	582	2 S42613	membrane protein p
34	73.5	11.5	438	2 A47702	glucan 1,3-beta-gl
35	73.5	11.5	438	2 T52149	beta-glucanase (im
36	73	11.4	861	2 S12499	CHLI protein - yea
37	72.5	11.3	386	2 S14570	tubulin beta chain
38	71.5	11.2	132	2 D88508	protein H14A12.5 f
39	71.5	11.2	1227	2 T23004	hypothetical prote
40	71.5	11.2	1779	2 T23130	hypothetical prote
41	71	11.1	677	1 SYECMT	methionine-tRNA li
42	71	11.1	677	2 C85839	methionine-tRNA sy
43	71	11.1	677	2 H90993	methionine tRNA sy
44	70.5	11.0	448	2 JN0118	glucan 1,3-beta-gl
45	70.5	11.0	464	2 F81278	probable periplasm

ALIGNMENTS

RESULT 1

A45361  
cystatin-related epididymal specific protein - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C;Accession: A45361  
R;Cornwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.  
Mol. Endocrinol. 6, 1653-1664, 1992  
A;Title: The CRSS gene: a unique testis-regulated gene related to the cystatin family is  
A;Reference number: A45361, MUID:93078799; PMID:1280328  
A;Accession: A45361  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: nucleic acid  
A;Residues: 1-139 <OR>  
A;Cross-references: GB:S49926; NID:G260492; PIDN:AAC35390.1; PID:G260493  
A;Note: sequence extracted from NCBI backbone (NCBIP:118813)  
C;Superfamily: cystatin; cystatin homology  
F;28-139/Domain: cystatin homology <Cys>

Query Match 33.4%; Score 214; DB 2; Length 139;  
Best Local Similarity 37.1%; Pred. No. 3.5e-14;  
Matches 43; Conservative 29; Mismatches 30; Indels 14; Gaps 4;

QY	12	SVHEWAVENY-----AKDSLQWTDQYNKESDDKYHFRIFRVLKQVQVTDHLS	61
DB	22	SKNEVKA-QNYFGSINISNANVKQVFWFAWKYNESEDKYFLVDKILHAKLQITDRME	80
QY	62	YHLNVEMQWTTCKP--ETTNCVPQER-ELHKQVNCFFSVFAVPWFPEQYKILNKSC	114
DB	81	YQIDVQISRNSCKKPLNNTENCIPQKKPLEKKNKSCFLVGALPWNGEFNLLSKEC	136

RESULT 2

S07085

cystatin C precursor - rat (fragment)

C;Species: Rattus norvegicus (Norway rat)  
C;Date: 01-Dec-1993 #sequence\_revision 03-Aug-1995 #text\_change 16-Jul-1999  
C;Accession: S07085; S01337; S21109  
R;Cole, T.; Dickson, P.W.; Esnard, F.; Averill, S.; Risbridger, G.P.; Gauthier, F.; Schre

Eur. J. Biochem. 186, 35-42, 1989.

A;Title: The cDNA structure and expression analysis of the genes for the cysteine protei

A;Reference number: S07085; MUID:90092122; PMID:2689174

A;Accession: S07085

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-127 <COL>

A;Cross-references: EMBL:X16957; NID:G56041; PIDN:CAA34831.1; PID:G736290

R;Esnard, A.; Esnard, F.; Faucher, D.; Gauthier, F.

FEBS Lett. 236, 475-478, 1988

A;Title: Two rat homologues of human cystatin C.

A;Reference number: S01337; MUID:88313020; PMID:3044831

A;Accession: S01337



; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
 ; FILE REFERENCE: 00-81PC  
 ; CURRENT APPLICATION NUMBER: US/09/941,314  
 ; CURRENT FILING DATE: 2001-08-29  
 ; PRIOR FILING DATE: 2001-09-01  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 36  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-941-314-6

Query Match 29.5%; Score 189; DB 9; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-14;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RKTFLSVHVMVENYAKDSLQWITDQYNKESDDK 41  
 Db 1 RKTFLSVHVMVENYAKDSLQWITDQYNKESDDK 36  
 RESULT 14  
 US-09-941-314-8  
 ; Sequence 8, Application US/09941314  
 ; Patent No. US20020142396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZymoGenetics, Inc.  
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
 ; FILE REFERENCE: 00-81PC  
 ; CURRENT APPLICATION NUMBER: US/09/941,314  
 ; CURRENT FILING DATE: 2001-08-29  
 ; PRIOR FILING DATE: 2001-09-01  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 35  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-941-314-8

Query Match 29.2%; Score 187; DB 9; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-14;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 KDSLQWITDQYNKESDDKXHFRIFRVLKVRQVTD 58  
 Db 1 KDSLQWITDQYNKESDDKXHFRIFRVLKVRQVTD 35  
 RESULT 15  
 US-09-740-638-2  
 ; Sequence 2, Application US/09740638  
 ; Patent No. US20020006656A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Holloway, James L.  
 ; TITLE OF INVENTION: Zcy85: A Member of the Cystatin  
 ; TITLE OF INVENTION: Superfamily  
 ; FILE REFERENCE: 99-104  
 ; CURRENT APPLICATION NUMBER: US/09/740,638  
 ; CURRENT FILING DATE: 2000-12-18  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: Fast-SEQ for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 145  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-740-638-2

Query Match 27.7%; Score 177.5; DB 9; Length 145;

Best Local Similarity 31.4%; Pred. No. 4.5e-12;  
 Matches 32; Conservative 30; Mismatches 37; Indels 3; Gaps 2;  
 QY 15 EVMAVENYAKDSLQWITDQYNKESDDKXHFRIFRVLKVRQVTDHLEYHLNVENQWITCQ 74  
 Db 34 KLMSKKN-MNSTLNFITQSYNNASNDTLYRVORLIRSORQLTGVGEYIVTVKIGWTKCK 92  
 QY 75 KPETTIN--CVPOREHLHKVNCFFSVFAVPWFQYKILNKSC 114  
 Db 93 RNDTSNSSCPLQSKKLKRLKSLICESLIYTMPEWYFQLWNNSC 134

Search completed: March 18, 2004, 14:30:52  
 Job time : 70.8036 secs

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RKKTFLSVHVMVAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 55  
 Db 1 RKKTFLSVHVMVAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 10  
 US-09-864-761-48936  
 ; Sequence 48936, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharon G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 ; FILE REFERENCE: Aeomica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 48936  
 ; LENGTH: 50  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AL096677.18  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96  
 ; OTHER INFORMATION: EST HUMAN HIT: AI200857.1, EVALU6 5.00e-23  
 ; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALU6 1.00e-01  
 US-09-864-761-48936

Query Match 40.9%; Score 262; DB 9; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RKKTFLSVHVMVAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 55  
 Db 1 RKKTFLSVHVMVAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 11  
 US-09-941-314-10  
 ; Sequence 10, Application US/09941314  
 ; Patent No. US20020142396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZymoGenetics, Inc.  
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
 ; FILE REFERENCE: 00-81PC  
 ; CURRENT APPLICATION NUMBER: US/09/941,314  
 ; CURRENT FILING DATE: 2001-08-29  
 ; PRIOR APPLICATION NUMBER: 60/230,230  
 ; PRIOR FILING DATE: 2001-09-01  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 10  
 ; LENGTH: 46  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-941-314-10

Query Match 39.7%; Score 254; DB 9; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 NKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCKPPTTN 80  
 Db 1 NKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCKPPTTN 46

RESULT 12  
 US-09-941-314-12  
 ; Sequence 12, Application US/09941314  
 ; Patent No. US20020142396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZymoGenetics, Inc.  
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
 ; FILE REFERENCE: 00-81PC  
 ; CURRENT APPLICATION NUMBER: US/09/941,314  
 ; CURRENT FILING DATE: 2001-08-29  
 ; PRIOR APPLICATION NUMBER: 60/230,230  
 ; PRIOR FILING DATE: 2001-09-01  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 12  
 ; LENGTH: 33  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-941-314-12

Query Match 29.5%; Score 189; DB 9; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-14;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 QVTDHLEYHLNVEMQWTTCKPPTTNVCVPQERE 87  
 Db 1 QVTDHLEYHLNVEMQWTTCKPPTTNVCVPQERE 33

RESULT 13  
 US-09-941-314-6  
 ; Sequence 6, Application US/09941314  
 ; Patent No. US20020142396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZymoGenetics, Inc.  
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to

```

; ORGANISM: Homo sapiens
US-09-941-314-14
Query Match 45.0%; Score 288; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 4.1e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 KESDDKXHYRIFRVLKQVQVTDHLEVLHNVNQWTTCKPPTNCVPOERE 87
DB 1 KESDDKXHYRIFRVLKQVQVTDHLEVLHNVNQWTTCKPPTNCVPOERE 52

RESULT 7
US-09-941-314-17
; Sequence 17, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-17

Query Match 42.7%; Score 273; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 NVEMQWTTCKPPTNCVPOERLHKQVNCFFSVFVAVPWFQYKILNK 112
DB 1 NVEMQWTTCKPPTNCVPOERLHKQVNCFFSVFVAVPWFQYKILNK 48

RESULT 8
US-09-941-314-13
; Sequence 13, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-13

Query Match 42.5%; Score 272; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.5e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 DQYNKESDDKXHYRIFRVLKQVQVTDHLEVLHNVNQWTTCKPPTN 80
DB 1 DQYNKESDDKXHYRIFRVLKQVQVTDHLEVLHNVNQWTTCKPPTN 49

RESULT 9
US-09-864-761-34822
; Sequence 34822, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34822
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109954.10
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
US-09-864-761-34822

Query Match 40.9%; Score 262; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.4e-22;

```

GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 137  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-2

Query Match 100.0%; Score 640; DB 9; Length 137;  
Best Local Similarity 100.0%; Pred. No. 2.2e-64;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPYQARKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHL 60  
DB 21 LPYQARKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHL 80

QY 61 EYHLNVMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFQYKILNKSCSSD 117  
DB 81 EYHLNVMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFQYKILNKSCSSD 137

RESULT 3  
US-09-941-314-3  
; Sequence 3, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-3

Query Match 98.3%; Score 629; DB 9; Length 115;  
Best Local Similarity 100.0%; Pred. No. 3e-63;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YQARKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEY 62  
DB 1 YQARKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEY 60

QY 63 HLNVMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFQYKILNKSCSSD 117  
DB 61 HLNVMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFQYKILNKSCSSD 115

RESULT 4  
US-09-941-314-15  
; Sequence 15, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC

GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-15

Query Match 69.7%; Score 446; DB 9; Length 80;  
Best Local Similarity 100.0%; Pred. No. 9.6e-43;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 QYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVMQWTTCKPETTNCVQRELHKQV 92  
DB 1 QYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVMQWTTCKPETTNCVQRELHKQV 60

QY 93 NCFPSVFAVPWFQYKILNK 112  
DB 61 NCFPSVFAVPWFQYKILNK 80

RESULT 5  
US-09-941-314-16  
; Sequence 16, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-16

Query Match 52.2%; Score 334; DB 9; Length 59;  
Best Local Similarity 100.0%; Pred. No. 3e-30;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 RQVTDHLEYHLNVMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFQYKILNK 112  
DB 1 RQVTDHLEYHLNVMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFQYKILNK 59

RESULT 6  
US-09-941-314-14  
; Sequence 14, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 52  
; TYPE: PRT

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 70.8036 Seconds  
(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-4  
Perfect score: 640  
Sequence: 1 LPQARKKTFSLVHEWAVE.....VFAVPWFQYKILNKSCSSD 117

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	640	100.0	117	9	US-09-941-314-4
2	640	100.0	137	9	US-09-941-314-2
3	629	98.3	115	9	US-09-941-314-3
4	446	69.7	80	9	US-09-941-314-15
5	334	52.2	59	9	US-09-941-314-16
6	288	45.0	52	9	US-09-941-314-14
7	273	42.7	48	9	US-09-941-314-17
8	272	42.5	49	9	US-09-941-314-13
9	262	40.9	50	9	US-09-864-761-34822
10	262	40.9	50	9	US-09-864-761-48936
11	254	39.7	46	9	US-09-941-314-10
12	189	29.5	33	9	US-09-941-314-12
13	189	29.5	36	9	US-09-941-314-6
14	187	29.2	35	9	US-09-941-314-8
15	177.5	27.7	145	9	US-09-740-638-2

16	177.5	27.7	145	13	US-10-006-467-2	Sequence 2, Appli
17	177.5	27.7	145	14	US-10-235-148-2	Sequence 2, Appli
18	171.5	26.8	127	8	US-08-849-303-19	Sequence 19, Appli
19	171.5	26.8	145	14	US-10-168-425-14	Sequence 14, Appli
20	170.5	26.6	116	9	US-09-775-932-16	Sequence 16, Appli
21	170.5	26.6	139	8	US-08-849-303-15	Sequence 15, Appli
22	170.5	26.6	139	9	US-09-969-834-4	Sequence 4, Appli
23	160.5	25.1	121	9	US-09-775-932-8	Sequence 8, Appli
24	160.5	25.1	140	14	US-10-376-564-46	Sequence 46, Appli
25	160.5	25.1	141	8	US-08-849-303-24	Sequence 24, Appli
26	160.5	25.1	141	9	US-09-940-497-6	Sequence 6, Appli
27	159.5	24.9	181	15	US-10-264-049-2608	Sequence 2608, Ap
28	157.5	24.6	165	9	US-09-740-638-5	Sequence 5, Appli
29	157.5	24.6	165	13	US-10-006-467-5	Sequence 5, Appli
30	157.5	24.6	165	14	US-10-235-148-5	Sequence 18, Appli
31	156.5	24.5	140	8	US-08-849-303-18	Sequence 2, Appli
32	156.5	24.5	140	14	US-10-376-564-48	Sequence 48, Appli
33	155	24.2	120	9	US-09-775-932-2	Sequence 17, Appli
34	155	24.2	146	8	US-08-849-303-17	Sequence 3, Appli
35	155	24.2	146	9	US-09-940-497-3	Sequence 3, Appli
36	155	24.2	146	9	US-09-969-834-3	Sequence 3, Appli
37	155	24.2	146	14	US-10-329-428-3	Sequence 3, Appli
38	155	24.2	146	14	US-10-376-564-47	Sequence 47, Appli
39	153.5	24.0	112	8	US-08-849-303-16	Sequence 16, Appli
40	153.5	24.0	118	9	US-09-775-932-24	Sequence 24, Appli
41	146.5	22.9	121	9	US-09-775-932-4	Sequence 4, Appli
42	146.5	22.9	141	8	US-08-849-303-22	Sequence 22, Appli
43	146.5	22.9	141	9	US-09-940-497-5	Sequence 5, Appli
44	146.5	22.9	141	9	US-09-974-298-141	Sequence 141, App
45	146.5	22.9	141	14	US-10-241-220-77	Sequence 77, Appli

## ALIGNMENTS

## RESULT 1

US-09-941-314-4  
; Sequence 4, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-4

Query Match 100.0%; Score 640; DB 9; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1.8e-64;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	LPQARKKTFSLVHEWAVENYAKDSIQWITDQYNKESDDKYHFRIPVLKQVQRTDHL	60
Db	1	LPQARKKTFSLVHEWAVENYAKDSIQWITDQYNKESDDKYHFRIPVLKQVQRTDHL	60
Qy	61	EYHLNEMQWTTQCKPCTTNCVPERELHKQVNCFFSVFVFPWFQYKILNKSCSSD	117
Db	61	EYHLNEMQWTTQCKPCTTNCVPERELHKQVNCFFSVFVFPWFQYKILNKSCSSD	117

## RESULT 2

US-09-941-314-2  
; Sequence 2, Application US/09941314  
; Patent No. US20020142396A1

```
; TELEFAX: 301 309 8512
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Cystatin SN
; US-08-744-138-6

Query Match      25.1%; Score 160.5; DB 3; Length 141;
Best Local Similarity 31.5%; Pred. No. 5.1e-12;
Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;

QY      26 SLOWITDQYNKESDDKYHFRIFRVLKVRQVTDHLEHYHLNVMQWTTCK--PETTNCVP 83
Db      47 ALHFVISEYNKATEDEYRRLRLRLRAREQIVGGVNYFFDIEVGRITCTKSPNLDTCF 106

QY      84 QER-ELHKQVNCFFSFVAFVWFQYKILNKSC 114
Db      107 HEQPELQKQKQCSFQIYEVPWEDRMSLVNSRC 138
```

Search completed: March 18, 2004, 14:25:34  
Job time : 29.9464 secs

Db 107 HDPEMAKYTTCTFVVVYSPWLNQIKLLESKC 138

RESULT 12

5432264-6

Patent No. 5432264

APPLICANT: GRUBB, ANDERS; LUNDWALL, AKE; ABRAHAMSON, MAGNUS;

DALBOGE, HENRIK

TITLE OF INVENTION: RECOMBINANT 3-DES-OH-CYSTATIN C PRODUCED

BY EXPRESSION IN A PROCARYOTIC HOST CELL

NUMBER OF SEQUENCES: 8

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/929,290

FILING DATE: 13-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 440,221

FILING DATE: 21-NOV-1989

APPLICATION NUMBER: 297,198

FILING DATE: 20-MAY-1988

SEQ ID NO: 6;

LENGTH: 146

5432264-6

Query Match 25.3%; Score 162; DB 6; Length 146;

Best Local Similarity 34.7%; Pred. No. 3.5e-12;

Matches 35; Conservative 21; Mismatches 41; Indels 4; Gaps 3;

QY 20 ENYAKDSLOWITDOYNKESDDKYHFRIVRLKQVQVTDHLEYHLNVEMQWTTCCQK--PE 77

Db 46 EGVRRALDPVAGCEYNKASNDYHSRALQVVRARKQIVAGVNYFLDVLGRITCTKTQPN 105

QY 78 TTNC-VPOERELHKQVNCFFSVFVAPVFPFQYKILNKSCSD 117

Db 106 LDCNCPHDPHKLKRAFCFQIYAVPW-QQTMTLSKSTCOD 145

RESULT 13

US-09-775-932-8

Sequence 8, Application US/09775932

Patent No. 6534477

GENERAL INFORMATION:

APPLICANT: University of British Columbia

TITLE OF INVENTION: Production and use of Modified Cystatins

FILE REFERENCE: 58069

CURRENT APPLICATION NUMBER: US/09/775,932

CURRENT FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: CA99/00717

PRIOR FILING DATE: 1999-08-05

PRIOR APPLICATION NUMBER: 60/095,503

PRIOR FILING DATE: 1998-08-05

NUMBER OF SEQ ID NOS: 32

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 8

LENGTH: 121

TYPE: PRT

ORGANISM: Homo sapiens

US-09-775-932-8

Query Match 25.1%; Score 160.5; DB 4; Length 121;

Best Local Similarity 31.5%; Pred. No. 4.2e-12;

Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDOYNKESDDKYHFRIVRLKQVQVTDHLEYHLNVEMQWTTCCQK--PETTNCVP 83

Db 27 ALHFVISEYNKATEDEYVRLRLVLRAREQIVGVVNYFFDIEVGRITCTKSQPNLDTCAF 86

QY 84 QER-ELHKQVNCFFSVFVAPVFPFQYKILNKSC 114

Db 87 HEQPELQKQLCSFQIYEVFVWEDRMSLVNSRC 118

RESULT 14

US-09-886-319A-46

Sequence 46, Application US/09886319A

Patent No. 6586185

GENERAL INFORMATION:

APPLICANT: Wolf, Eckard

APPLICANT: Werner, Sabine

APPLICANT: Halle, Jorn-Peter

APPLICANT: Regenbogen, Johannes

APPLICANT: Goppelt, Andreas

TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for

TITLE OF INVENTION: the diagnosis or treatment of Skin Disorders and wound

TITLE OF INVENTION: Healing and for the Identification of Pharmacologically

TITLE OF INVENTION: Active Substances

FILE REFERENCE: 50125/014002

CURRENT APPLICATION NUMBER: US/09/886,319A

CURRENT FILING DATE: 2001-06-20

PRIOR APPLICATION NUMBER: US 60/222,081

PRIOR FILING DATE: 2000-08-01

PRIOR APPLICATION NUMBER: DE 10030149.5

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 84

SOFTWARE: PastSeq for Windows Version 4.0

SEQ ID NO 46

LENGTH: 140

TYPE: PRT

ORGANISM: Mus musculus

US-09-886-319A-46

Query Match 25.1%; Score 160.5; DB 4; Length 140;

Best Local Similarity 32.4%; Pred. No. 5.1e-12;

Matches 34; Conservative 24; Mismatches 44; Indels 3; Gaps 2;

QY 15 EYMAVENYAKDSLOWITDOYNKESDDKYHFRIVRLKQVQVTDHLEYHLNVEMQWTTCCQ 74

Db 35 EADANEEGVRRALDPVAGVSEYNGSDYHSRALQVVRARKQIVAGVNYFLDVLGRITCT 94

QY 75 KPET--TNC-VPOERELHKQVNCFFSVFVAPVFPFQYKILNKSCSS 116

Db 95 KSQTNLTDCPFHDPHKLKALCSFQIYSPVPMKGTSLTKFSCKN 139

RESULT 15

US-08-744-138-6

Sequence 6, Application US/08744138

Patent No. 6011012

GENERAL INFORMATION:

APPLICANT: Gentz, Reiner L.

APPLICANT: Ni, Jian

APPLICANT: Rosen, Craig A.

APPLICANT: Yu, Guo-Liang

TITLE OF INVENTION: Human Cystatin E

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/744,138

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PF202P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301 309 8504

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0193 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 118195
;
US-08-791-522-4
Query Match 26.6%; Score 170.5; DB 2; Length 139;
Best Local Similarity 37.0%; Pred. No. 3e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDQYNKSDDKYHFRIFRLVKVQRQVTDHLEYHLNVEMQWTTCKP--ETNC-V 82
Db 47 ALOFAMAENRASNDKYSRRVVRVISAKROLVSGIKYILQVEIGRTTCPKSSGDLQSCF 106

QY 83 POERELHKQVNCFFSVFAVPWFQYKILNKSC 114
Db 107 HDEPEMAKYTTCTFVVVYSIPWLNQIKLLESKC 138

RESULT 10
US-09-314-777-4
; Sequence 4, Application US/09314777
; Patent No. 6110686
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Suzya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/314,777
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,522
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0193 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 118195
;
US-09-314-777-4
Query Match 26.6%; Score 170.5; DB 3; Length 139;
Best Local Similarity 37.0%; Pred. No. 3e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDQYNKSDDKYHFRIFRLVKVQRQVTDHLEYHLNVEMQWTTCKP--ETNC-V 82
Db 47 ALOFAMAENRASNDKYSRRVVRVISAKROLVSGIKYILQVEIGRTTCPKSSGDLQSCF 106

QY 83 POERELHKQVNCFFSVFAVPWFQYKILNKSC 114
Db 107 HDEPEMAKYTTCTFVVVYSIPWLNQIKLLESKC 138

RESULT 11
US-08-849-303-15
; Sequence 15, Application US/08849303
; Patent No. 6680424
; GENERAL INFORMATION:
; APPLICANT: Atkinson, Howard J.
; APPLICANT: McPherson, Michael J.
; APPLICANT: Urwin, Peter E.
; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 21-MAY-1997
; APPLICATION NUMBER: US/08/849,303
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
;
US-08-849-303-15
Query Match 26.6%; Score 170.5; DB 4; Length 139;
Best Local Similarity 37.0%; Pred. No. 3e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDQYNKSDDKYHFRIFRLVKVQRQVTDHLEYHLNVEMQWTTCKP--ETNC-V 82
Db 47 ALOFAMAENRASNDKYSRRVVRVISAKROLVSGIKYILQVEIGRTTCPKSSGDLQSCF 106

QY 83 POERELHKQVNCFFSVFAVPWFQYKILNKSC 114

```

```

QY 15 EWMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVQROVTDHLEYHLNVMQWTTCCQ 7
Db 22 EDADSEEGVQRALDFAVSEYNGKSGNDAYSRAIQVVRARKQLVAGIYLDVEMGRITTC 81
QY 75 KPET--TNC-VQERELHKQVNCFFSFVAFWPFEOYKILNKSCSS 116
Db 82 KSQTNLTNCFPHDQPHLMRKALCSFQIYSVPWKGTHTLTKSCKN 126

RESULT 8
US-09-775-932-16
; Sequence 16, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cystatins
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 116
; TYPE: PRN
; ORGANISM: Gallus sp.
US-09-775-932-16

Query Match 26.6%; Score 170.5; DB 4; Length 116;
Best Local Similarity 37.0%; Pred. No. 2.3e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 3

QY 26 SLOWITDQYNKESDDKYHFRIFRVLKVQROVTDHLEYHLNVMQWTTCCQK 82
Db 24 ALQFMAEYTRASNQKYSRVRVISAARQLVSGIKYLOVEIGRTTCPSKSGDLOSCEF 83
QY 83 PQERELHKQVNCFFSFVAFWPFEOYKILNKSC 114
Db 84 HDEPEMAKYTCTCFVYVSIPMLNQIKLESKC 115

RESULT 9
US-08-791-522-4
; Sequence 4, Application US/08791522
; Patent No. 5935817
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,522
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

; PRIOR FILING DATE: 1998-11-20  
; PRIOR APPLICATION NUMBER: 60/156,382  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 142  
; TYPE: PR1  
; ORGANISM: Mus musculus  
US-09-617-302-3

Query Match 33.4%; Score 214; DB 3; Length 142;  
Best Local Similarity 37.4%; Pred. No. 1.4e-18;  
Matches 43; Conservative 29; Mismatches 30; Indels 14; Gaps 4;  
QY 12 SVHEVMAVENY-----AKDSLOWITDOYNKESDDKYHFRIFRVLKVRQVTDHLE 61  
DB 25 SKNEVA-QNYFGSINISNANVCQWPFAMKEYNKESDKYFLVDKILHAKLQITDRME 83  
QY 62 YHLNVMQWTTCKP--ETTNCVPQER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 114  
DB 84 YQIDVOISRSNCKPLNNTENCIPQKPELEKMKSCFLVGALPWNGEFNLKSC 139

## RESULT 3

US-09-431-480-2  
; Sequence 2, Application US/09431480  
; Patent No. 6235708  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72  
; CURRENT APPLICATION NUMBER: US/09/431,480  
; CURRENT FILING DATE: 1999-11-01  
; EARLIER APPLICATION NUMBER: 60/109,217  
; EARLIER FILING DATE: 1998-11-20  
; EARLIER APPLICATION NUMBER: 60/156,382  
; EARLIER FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 141  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-431-480-2

Query Match 29.5%; Score 188.5; DB 3; Length 141;  
Best Local Similarity 34.5%; Pred. No. 1.9e-15;  
Matches 40; Conservative 25; Mismatches 38; Indels 13; Gaps 3;  
QY 2 PYQARKTFLSVHEVMAVENYAKDSLOWITDOYNKESDDKYHFRIFRVLKVRQVTDHLE 61  
DB 33 PFEDIPKSYV-----YVQHALWYAMKEYNKASNDLYNFRVVDILKSQEQITDSLE 82  
QY 62 YHLNVMQWTTCKP--ETTNCVPQER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 114  
DB 83 YYLEVNIARTMCKKIAGDNENCLFQODPKMKMWFCIFIVSSKPKWPKELKMLKKQC 138

## RESULT 4

US-09-617-302-2  
; Sequence 2, Application US/09617302  
; Patent No. 6245529  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72 C1  
; CURRENT APPLICATION NUMBER: US/09/617,302  
; CURRENT FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/431,480  
; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: 60/109,217  
; PRIOR FILING DATE: 1998-11-20  
; PRIOR APPLICATION NUMBER: 60/156,382  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 141  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-617-302-2

Query Match 29.5%; Score 188.5; DB 3; Length 141;  
Best Local Similarity 34.5%; Pred. No. 1.9e-15;  
Matches 40; Conservative 25; Mismatches 38; Indels 13; Gaps 3;  
QY 2 PYQARKTFLSVHEVMAVENYAKDSLOWITDOYNKESDDKYHFRIFRVLKVRQVTDHLE 61  
DB 33 PFEDIPKSYV-----YVQHALWYAMKEYNKASNDLYNFRVVDILKSQEQITDSLE 82  
QY 62 YHLNVMQWTTCKP--ETTNCVPQER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 114  
DB 83 YYLEVNIARTMCKKIAGDNENCLFQODPKMKMWFCIFIVSSKPKWPKELKMLKKQC 138

## RESULT 5

US-09-431-480-4  
; Sequence 4, Application US/09431480  
; Patent No. 6235708  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72  
; CURRENT APPLICATION NUMBER: US/09/431,480  
; CURRENT FILING DATE: 1999-11-01  
; EARLIER APPLICATION NUMBER: 60/109,217  
; EARLIER FILING DATE: 1998-11-20  
; EARLIER APPLICATION NUMBER: 60/156,382  
; EARLIER FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 142  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-431-480-4

Query Match 29.5%; Score 188.5; DB 3; Length 142;  
Best Local Similarity 40.4%; Pred. No. 1.9e-15;  
Matches 38; Conservative 23; Mismatches 30; Indels 3; Gaps 2;  
QY 24 KDSLOWITDOYNKESDDKYHFRIFRVLKVRQVTDHLEHLNVMQWTTCKPPTTN--C 81  
DB 46 KQCLWFAMQYKNESEDKYFLVVKTLQALQVTLNLEVLIDVEIARSDCKPLSTNEIC 105  
QY 82 VPOER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 114  
DB 106 AIQENSKKLKLSCSFLVGALPWNGEFTVMEKCC 139

## RESULT 6

US-09-617-302-4  
; Sequence 4, Application US/09617302  
; Patent No. 6245529  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72 C1  
; CURRENT APPLICATION NUMBER: US/09/617,302  
; CURRENT FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/431,480

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 29.9464 Seconds  
(without alignments)  
201.701 Million cell updates/sec

Title: US-09-941-314-4  
Perfect score: 640  
Sequence: 1 LPQARKTKFLSVHEWAVE.....VFAVPWFQYKILNKSCSD 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pap.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pap.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pap.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pap.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pap.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214	33.4	142	3	US-09-431-480-3
2	214	33.4	142	3	US-09-617-302-3
3	188.5	29.5	141	3	US-09-431-480-2
4	188.5	29.5	141	3	US-09-617-302-2
5	188.5	29.5	142	3	US-09-431-480-4
6	188.5	29.5	142	3	US-09-617-302-4
7	171.5	26.8	127	4	US-08-849-303-19
8	170.5	26.6	116	4	US-09-775-932-16
9	170.5	26.6	139	2	US-08-791-522-4
10	170.5	26.6	139	3	US-09-314-777-4
11	170.5	26.6	139	4	US-08-849-303-15
12	162	25.3	146	6	5432264-6
13	160.5	25.1	121	4	US-09-775-932-8
14	160.5	25.1	140	4	US-09-886-319A-46
15	160.5	25.1	141	3	US-08-744-138-6
16	160.5	25.1	141	4	US-09-241-376-6
17	160.5	25.1	141	4	US-09-940-497-6
18	160.5	25.1	141	4	US-08-849-303-24
19	156.5	24.5	140	3	US-09-431-480-5
20	156.5	24.5	140	3	US-09-617-302-5
21	156.5	24.5	140	4	US-09-886-319A-48
22	156.5	24.5	140	4	US-08-849-303-18
23	155	24.2	120	6	5432264-4
24	155	24.2	145	2	US-08-832-535-11
25	155	24.2	146	2	US-08-791-522-3
26	155	24.2	146	3	US-08-744-138-3
27	155	24.2	146	3	US-08-744-138-3

28	155	24.2	146	3	US-09-019-485-4	Sequence 4, Appli
29	155	24.2	146	3	US-09-314-777-3	Sequence 3, Appli
30	155	24.2	146	3	US-09-431-480-6	Sequence 6, Appli
31	155	24.2	146	3	US-09-617-302-6	Sequence 6, Appli
32	155	24.2	146	4	US-09-241-376-3	Sequence 3, Appli
33	155	24.2	146	4	US-09-528-436B-3	Sequence 3, Appli
34	155	24.2	146	4	US-09-886-319A-47	Sequence 47, Appli
35	155	24.2	146	4	US-09-940-497-3	Sequence 3, Appli
36	155	24.2	146	4	US-09-976-594-37	Sequence 17, Appli
37	155	24.2	146	4	US-08-849-303-17	Sequence 17, Appli
38	155	24.2	146	5	PCT-US95-07135-9	Sequence 9, Appli
39	153.5	24.0	112	4	US-08-849-303-16	Sequence 16, Appli
40	153.5	24.0	118	4	US-09-775-932-24	Sequence 24, Appli
41	146.5	22.9	121	4	US-09-775-932-4	Sequence 4, Appli
42	146.5	22.9	141	3	US-08-744-138-5	Sequence 5, Appli
43	146.5	22.9	141	3	US-09-431-480-11	Sequence 11, Appli
44	146.5	22.9	141	3	US-09-617-302-11	Sequence 11, Appli
45	146.5	22.9	141	4	US-09-241-376-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-09-431-480-3  
; Sequence 3, Application US/09431480  
; Patent No. 6235708  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72  
; CURRENT APPLICATION NUMBER: US/09/431,480  
; CURRENT FILING DATE: 1999-11-01  
; EARLIER APPLICATION NUMBER: 60/109,217  
; EARLIER FILING DATE: 1998-11-20  
; EARLIER APPLICATION NUMBER: 60/156,382  
; EARLIER FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-431-480-3

Query Match	33.4%	Score 214;	DB 3;	Length 142;
Best Local Similarity	37.1%	Pred. No. 1.4e-18;		
Matches	43;	Conservative	29;	Mismatches 30; Indels 14; Gaps 4;
QY	12	SVHEWAVENY-----AKDSLOWITDQNKESDDKYHFRIFRVLKVGQVTDHLE	61	
DB	25	SKNEVKA-QNYFGSINISNANVKQCVFAMKEYKESDKYVFLVDKILHAKLQITDRME	83	
QY	62	YHLNVEWQMTTCKP--ETTNCVQPER-ELHKVNCFFSVFAVPWFQYKILNKSC	114	
DB	84	YQIDVOISRNCKPLNNTENCIPQKKPELEKQKSCFLVGALPWNGEFLLSKEC	139	

RESULT 2  
US-09-617-302-3  
; Sequence 3, Application US/09617302  
; Patent No. 6245529  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72 C1  
; CURRENT APPLICATION NUMBER: US/09/617,302  
; CURRENT FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/431,480  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: 60/109,217

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488899/53.  
 DR Single exon nucleic acid probes for analyzing gene expression in human  
 XX hearts.  
 PT Claim 15; SEQ ID NO 21294; 530pp; English.  
 PS  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41105). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 50 AA;

Query Match 40.9%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVRQ 55  
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVRQ 50

RESULT 15  
 AM67252  
 ID AM67252 standard; protein; 50 AA.  
 XX  
 AC AM67252;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27558.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488900/53.  
 XX  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.  
 XX  
 XX Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.  
 PS  
 XX

CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention  
 XX  
 SQ Sequence 50 AA;  
 Query Match 40.9%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVRQ 55  
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVRQ 50

Search completed: March 18, 2004, 14:14:56  
 Job time : 106.089 secs

XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX PI WPI; 2001-496933/54.  
 XX DR  
 XX PT New spatially-addressable set of single exon nucleic acid probes, useful  
 XX PT for measuring gene expression in sample derived from human breast,  
 XX PT comprises number of single exon nucleic acid probes.  
 XX PS Claim 27; SEQ ID NO 15357; 327pp + Sequence Listing; English.  
 XX PS  
 XX CC The invention relates to a spatially-addressable set of single exon  
 XX CC nucleic acid probes for measuring gene expression in a sample derived  
 XX CC from human breast and BT 474 cells. The method involves contacting the  
 XX CC probes with a collection of detectably labelled nucleic acids derived  
 XX CC from mRNA of human breast, and then measuring the label bound to each  
 XX CC probe of the microarray. The probes are useful for verifying the  
 XX CC expression of regions of genomic DNA predicted to encode proteins. They  
 XX CC are useful for gene discovery, and for determining predisposition and/or  
 XX CC assessing the toxicity of chemical agents on cells. The microarray of  
 XX CC this invention presents a far greater diversity of probes for measuring  
 XX CC gene expression, with far less bias than expressed sequence tag  
 XX CC microarrays. The method is suitable for rapid production of functional  
 XX CC information from genomic sequence. The present sequence is a peptide  
 XX CC encoded by a single exon nucleic acid probe of the invention. Note: The  
 XX CC sequence data for this patent did not form part of the printed  
 XX CC specification, but was obtained in electronic format directly from WIPO  
 XX CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 50 AA;  
 Query Match 40.9%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 6 RKKTFLSVHVWAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQRQ 55  
 DB 1 RKKTFLSVHVWAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQRQ 50  
 RESULT 13  
 ABB28913  
 ID ABB28913 standard; peptide; 50 AA.  
 XX AC ABB28913;  
 XX DT 01-FEB-2002 (first entry)  
 XX DE Peptide #1564 encoded by breast cell single exon nucleic acid probe.  
 XX KW Human; microarray; single exon probe; gene expression; breast; disease;  
 XX KW cancer.  
 XX OS Homo sapiens.  
 XX PN WO200157271-A2.  
 XX PD 09-AUG-2001.  
 XX PF 30-JAN-2001; 2001WO-US0000662.  
 XX PR 04-FEB-2000; 2000US-0180312P.  
 XX PR 26-MAY-2000; 2000US-0207456P.  
 XX PR 30-JUN-2000; 2000US-00608408.  
 XX PR 03-AUG-2000; 2000US-00632366.  
 XX PR 21-SEP-2000; 2000US-0234687P.  
 XX PR 27-SEP-2000; 2000US-0236359P.  
 XX PR 04-OCT-2000; 2000GB-00024263.  
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-496933/54.  
 XX PT New spatially-addressable set of single exon nucleic acid probes, useful  
 XX PT for measuring gene expression in sample derived from human breast,  
 XX PT comprises number of single exon nucleic acid probes.  
 XX PS Claim 27; SEQ ID NO 11881; 327pp + Sequence Listing; English.  
 XX PS  
 XX CC The invention relates to a spatially-addressable set of single exon  
 XX CC nucleic acid probes for measuring gene expression in a sample derived  
 XX CC from human breast and BT 474 cells. The method involves contacting the  
 XX CC probes with a collection of detectably labelled nucleic acids derived  
 XX CC from mRNA of human breast, and then measuring the label bound to each  
 XX CC probe of the microarray. The probes are useful for verifying the  
 XX CC expression of regions of genomic DNA predicted to encode proteins. They  
 XX CC are useful for gene discovery, and for determining predisposition and/or  
 XX CC assessing the toxicity of chemical agents on cells. The microarray of  
 XX CC this invention presents a far greater diversity of probes for measuring  
 XX CC gene expression, with far less bias than expressed sequence tag  
 XX CC microarrays. The method is suitable for rapid production of functional  
 XX CC information from genomic sequence. The present sequence is a peptide  
 XX CC encoded by a single exon nucleic acid probe of the invention. Note: The  
 XX CC sequence data for this patent did not form part of the printed  
 XX CC specification, but was obtained in electronic format directly from WIPO  
 XX CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 50 AA;  
 Query Match 40.9%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 6 RKKTFLSVHVWAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQRQ 55  
 DB 1 RKKTFLSVHVWAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQRQ 50  
 RESULT 14  
 ABB19524  
 ID ABB19524 standard; protein; 50 AA.  
 XX AC ABB19524;  
 XX DT 23-JAN-2002 (first entry)  
 XX DE Protein #1523 encoded by probe for measuring heart cell gene expression.  
 XX KW Human; gene expression; heart; microarray; vascular system;  
 XX KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 XX KW congenital heart disease.  
 XX OS Homo sapiens.  
 XX PN WO200157274-A2.  
 XX PD 09-AUG-2001.  
 XX PF 30-JAN-2001; 2001WO-US0000666.  
 XX PR 04-FEB-2000; 2000US-0180312P.  
 XX PR 26-MAY-2000; 2000US-0207456P.  
 XX PR 30-JUN-2000; 2000US-00608408.  
 XX PR 03-AUG-2000; 2000US-00632366.  
 XX PR 21-SEP-2000; 2000US-0234687P.  
 XX PR 27-SEP-2000; 2000US-0236359P.  
 XX PR 04-OCT-2000; 2000GB-00024263.  
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.

```
RESULT 10
ABB34086
ID ABB34086 standard; peptide; 50 AA.
XX
AC ABB34086;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #1592 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-483447/52.
XX
Human genome-derived single exon nucleic acid probes useful for analyzing
gene expression in human fetal liver.
XX
Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.
XX
The invention relates to a single exon nucleic acid probe for measuring
human gene expression in a sample derived from human foetal liver. The
single exon nucleic acid probes may be used for predicting, measuring and
displaying gene expression in samples derived from human fetal liver. The
present sequence is a peptide encoded by a single exon nucleic acid probe
of the invention. Note: The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 50 AA;
Query Match 40.9%; Score 262; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 55
Db 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50
RESULT 11
AAM27545
ID AAM27545 standard; protein; 50 AA.
XX
AC AAM27545;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #1582 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX
genetic disorder.
XX
```

```
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488897/53.
XX
Human genome-derived single exon nucleic acid probes useful for analyzing
gene expression in human placenta.
XX
Claim 27; SEQ ID NO 27814; 654pp; English.
XX
The present invention relates to single exon nucleic acid probes (SENP:
see AAI31315-AAI57546). The present sequence is a peptide encoded by one
such probe. The probes are useful for producing a microarray for
predicting, measuring and displaying gene expression in samples derived
from human placenta. The probes are useful for antenatal diagnosis of
human genetic disorders
XX
SQ Sequence 50 AA;
Query Match 40.9%; Score 262; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 55
Db 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50
RESULT 12
ABB32389
ID ABB32389 standard; peptide; 50 AA.
XX
AC ABB32389;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #5040 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
```

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX  
 SQ Sequence 48 AA;  
 Query Match 42.7%; Score 273; DB 5; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-23;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 65 NVEMQWTTCKPPTNCVQPERLHKQVNCFFSVFAVPWFQYKILNK 112  
 Db 1 NVEMQWTTCKPPTNCVQPERLHKQVNCFFSVFAVPWFQYKILNK 48  
 RESULT 8  
 AAU79863  
 ID AAU79863 standard; peptide; 49 AA.  
 XX  
 AC AAU79863;  
 XX  
 DT 15-JUL-2002 (first entry)  
 DE Human cystatin-8 (Zcys8) antigenic fragment #11.  
 XX  
 DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 23-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230320P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 97-98; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX  
 SQ Sequence 49 AA;  
 Query Match 42.5%; Score 272; DB 5; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-23;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 32 DOYNKESDDKYHFRIFRVLKVRQVTDHLEYHLNVEMQWTTCKPPTN 80  
 Db 1 DOYNKESDDKYHFRIFRVLKVRQVTDHLEYHLNVEMQWTTCKPPTN 49  
 RESULT 9  
 AAM15096  
 ID AAM15096 standard; protein; 50 AA.  
 XX  
 AC AAM15096;  
 XX  
 DT 12-OCT-2001 (first entry)  
 DE Peptide #1530 encoded by probe for measuring cervical gene expression.  
 XX  
 DE Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-488901/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human cervical epithelial cells.  
 XX  
 PS Claim 27; SEQ ID NO 19922; 487pp; English.  
 XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 50 AA;  
 Query Match 40.9%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 RKKTFUSVHEWMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVRQ 55  
 Db 1 RKKTFUSVHEWMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVRQ 50

CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
SQ Sequence 59 AA;  
Query Match 52.2%; Score 334; DB 5; Length 59;  
Best Local Similarity 100.0%; Pred. No. 2.2e-30; Indels 0; Gaps 0;  
Matches 59; Conservative 0; Mismatches 0;  
QY 54 RQVTDHLEYHLNVEMQWTTCKPQREHLKQVNCFFSFVAFWPFQYKILNK 112  
DB 1 RQVTDHLEYHLNVEMQWTTCKPQREHLKQVNCFFSFVAFWPFQYKILNK 59  
RESULT 6  
AAU79864  
ID AAU79864 standard; peptide; 52 AA.  
XX  
AC AAU79864;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human cystatin-8 (Zcys8) antigenic fragment #12.  
XX  
KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic peptide.  
XX  
OS Homo sapiens.  
XX  
PN WO200220567-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US026868.  
XX  
PR 01-SEP-2000; 2000US-0230230P.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Holloway JL, Gao Z, Bishop PD;  
XX  
DR WPI; 2002-383044/41.  
XX  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX  
PS Claim 2; Page 98; 100pp; English.  
XX  
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
SQ Sequence 59 AA;

CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
SQ Sequence 52 AA;  
Query Match 45.0%; Score 288; DB 5; Length 52;  
Best Local Similarity 100.0%; Pred. No. 3.3e-25;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 36 KESDDKYHFRIPRLVKVQRQVTDHLEYHLNVEMQWTTCKPQREHLKQVNCFFSFVAFWPFQYKILNK 87  
DB 1 KESDDKYHFRIPRLVKVQRQVTDHLEYHLNVEMQWTTCKPQREHLKQVNCFFSFVAFWPFQYKILNK 52  
RESULT 7  
AAU79867  
ID AAU79867 standard; peptide; 48 AA.  
XX  
AC AAU79867;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human cystatin-8 (Zcys8) antigenic fragment #15.  
XX  
KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic peptide.  
XX  
OS Homo sapiens.  
XX  
PN WO200220567-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US026868.  
XX  
PR 01-SEP-2000; 2000US-0230230P.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Holloway JL, Gao Z, Bishop PD;  
XX  
DR WPI; 2002-383044/41.  
XX  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX  
PS Claim 2; Page 99; 100pp; English.  
XX  
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 XX  
 SQ Sequence 115 AA;

Query Match 98.3%; Score 629; DB 5; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-63;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 YQARKKTFSLVHVMVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 62  
 Db 1 YQARKKTFSLVHVMVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 60  
 Qy 63 HLNVENQWTTTCQKPEITNCVPQRELHKQVNCFFSVFAVPWFQYKILNKSCSSD 117  
 Db 61 HLNVENQWTTTCQKPEITNCVPQRELHKQVNCFFSVFAVPWFQYKILNKSCSSD 115

RESULT 4  
 AAU79865  
 ID AAU79865 standard; peptide; 80 AA.  
 AC AAU79865;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #13.  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 98; 100pp; English.

XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in

CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX  
 SQ Sequence 80 AA;

Query Match 69.7%; Score 446; DB 5; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-43;  
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 QYNKESDDKYHFRIFRLVKVQROVTDHLEYHLNVENQWTTTCQKPEITNCVPQRELHKQV 92  
 Db 1 QYNKESDDKYHFRIFRLVKVQROVTDHLEYHLNVENQWTTTCQKPEITNCVPQRELHKQV 60  
 Qy 93 NCFFSVFAVPWFQYKILNK 112  
 Db 61 NCFFSVFAVPWFQYKILNK 80

RESULT 5  
 AAU79866  
 ID AAU79866 standard; peptide; 59 AA.  
 AC AAU79866;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #14.  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 99; 100pp; English.

XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 CC  
 XX  
 SQ Sequence 117 AA;

Query Match 100.0%; Score 640; DB 5; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 7e-65;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDOYNKESDDKYHFRIFRVLKQVQVTDHL 60  
 DB 1 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDOYNKESDDKYHFRIFRVLKQVQVTDHL 60  
 QY 61 EYHLNEMQWTTCKPETTNCVPORELHKQVNCFFSVFAPWPFQYKILNKSCSSD 117  
 DB 61 EYHLNEMQWTTCKPETTNCVPORELHKQVNCFFSVFAPWPFQYKILNKSCSSD 117

RESULT 2  
 AAU79852  
 ID AAU79852 standard; protein; 137 AA.  
 XX  
 AC AAU79852;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8).  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 DR N-PSDB; ABK49522.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 93-94; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 137 AA;

Query Match 100.0%; Score 640; DB 5; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-65;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDOYNKESDDKYHFRIFRVLKQVQVTDHL 60  
 DB 21 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDOYNKESDDKYHFRIFRVLKQVQVTDHL 80  
 QY 61 EYHLNEMQWTTCKPETTNCVPORELHKQVNCFFSVFAPWPFQYKILNKSCSSD 117  
 DB 81 EYHLNEMQWTTCKPETTNCVPORELHKQVNCFFSVFAPWPFQYKILNKSCSSD 137

RESULT 3  
 AAU79853  
 ID AAU79853 standard; protein; 115 AA.  
 XX  
 AC AAU79853;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #1.  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic fragment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 DR  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 94; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 106.089 Seconds  
(without alignments)  
311.606 Million cell updates/sec

Title: US-09-941-314-4  
Perfect score: 640  
Sequence: 1 LPQARKKTFSLVHEWAVE.....VFAVPWFQYKILNKSCSSD 117

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:.\*  
1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001s:.\*  
5: Geneseq2002s:.\*  
6: Geneseq2003as:.\*  
7: Geneseq2003bs:.\*  
8: Geneseq2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	640	100.0	117	5	AAU79854	Human cys
2	640	100.0	137	5	AAU79852	Human cys
3	629	98.3	115	5	AAU79853	Human cys
4	446	69.7	80	5	AAU79865	Human cys
5	334	52.2	59	5	AAU79866	Human cys
6	288	45.0	52	5	AAU79864	Human cys
7	273	42.7	48	5	AAU79867	Human cys
8	272	42.5	49	5	AAU79863	Human cys
9	262	40.9	50	4	AAU15096	Peptide #
10	262	40.9	50	4	ABB34086	Peptide #
11	262	40.9	50	4	AAU72545	Peptide #
12	262	40.9	50	4	ABB32389	Peptide #
13	262	40.9	50	4	ABB28913	Peptide #
14	262	40.9	50	4	ABB19524	Protein #
15	262	40.9	50	4	AAU67252	Human bon
16	262	40.9	50	4	AAU54871	Human bon
17	262	40.9	50	4	ABGA48915	Human liv
18	262	40.9	50	4	AAU02833	Peptide #
19	262	40.9	50	5	ABG36903	Human pep
20	254	39.7	46	5	AAU79860	Human cys
21	214	33.4	142	4	AAE02404	Murine cy
22	214	33.4	142	4	AAE04433	Mouse cys
23	214	33.4	143	6	ADA14374	Mouse spe
24	199	31.1	142	7	ADD46708	Rat Prote
25	199	31.1	142	7	ADD46704	Rat Prote

## ALIGNMENTS

RESULT 1  
AAU79854  
ID AAU79854 standard; protein; 117 AA.

XX AAU79854;  
XX AC  
XX 15-JUL-2002 (first entry)  
XX DE Human cystatin-8 (zcys8) antigenic fragment #2.  
XX KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic fragment.  
XX OS Homo sapiens.  
XX PN WO200220567-A2.  
XX PD 14-MAR-2002.  
XX PF 29-AUG-2001; 2001WO-US026868.  
XX PR 01-SEP-2000; 2000US-0230230P.  
XX (ZYMO ) ZYMOGENETICS INC.  
XX Holloway JL, Gao Z, Bishop PD;  
WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.

Claim 2; Page 94-95; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8) polypeptide (I). (I) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcys8 is useful as research reagent for characterising sites of interaction between Zcys8 and its receptor. Zcys8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(I) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcys8. The antibodies are also useful to isolate large

AAU79862 Human cys  
AAU79856 Human cys  
AAU78260 Fragment  
AAU96576 Murine cy  
AAU04403 Murine cy  
AAU04432 Mouse tes  
AAU78258 Fragment  
AAU02405 Human cys  
AAU04434 Human cys  
ADA57231 Human sec  
ADA41112 Human sec  
ADC74335 Human sec  
ADC7980 Human sec  
ADD46706 Human Pro  
ADD46710 Human Pro  
AAU79858 Human cys  
AAU78259 Fragment  
AAU04315 Alternati  
AAU76555 Human Zcy  
ABG75917 Human cys

26 189 29.5 33 5 AAU79862  
27 189 29.5 36 5 AAU79856  
28 188.5 29.5 123 5 AAU78260  
29 188.5 29.5 141 3 AAU96576  
30 188.5 29.5 141 4 AAU02403  
31 188.5 29.5 141 4 AAU04432  
32 188.5 29.5 142 2 AAU78258  
33 188.5 29.5 142 4 AAU02405  
34 188.5 29.5 142 4 AAU04434  
35 188.5 29.5 142 6 ADA57231  
36 188.5 29.5 142 6 ADA41112  
37 188.5 29.5 142 7 ADC74335  
38 188.5 29.5 142 7 ADC7980  
39 188.5 29.5 142 7 ADD46706  
40 188.5 29.5 142 7 ADD46710  
41 187 29.2 35 5 AAU79858  
42 184.5 28.8 92 2 AAU78259  
43 177.5 27.7 145 4 AAU04315  
44 177.5 27.7 145 5 AAU76555  
45 177.5 27.7 145 6 ABG75917

Search completed: March 18, 2004, 14:21:23  
Job time : 70.9087 secs

```
RA Ikegawa S., Nakamura Y.;
RT "DD72, a novel mouse gene implicated in the early stage of ectopic
RL ossification.";
RN Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036743; BAA95411.1; -
DR EMBL; BC048364; AAH48364.1; -
DR HSSP; P01034; IG96.
DR MGD; MGI:193004; Cst10.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR001713; Stefina.
DR Pfam; PF00031; cystatin; 1.
DR PRINTS; PR00295; STEFINA.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
SQ SEQUENCE 148 AA; 16451 MW; 637534CBFCSAA179 CRC64;

Query Match 18.0%; Score 113; DB 11; Length 148;
Best Local Similarity 28.2%; Pred. No. 0.00024;
Matches 29; Conservative 18; Mismatches 46; Indels 10; Gaps 3;

QY 13 EVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVRQVTDHLEYHLNVEMQWTTCC 72
DB 50 EVQKVVEA-----VRTYNDNDNDLYLSKPIRLMSASQQVVAGKYYLYKIELGRTTCT 102

QY 73 KPET--TNCVQPQR-ELHKQVNCFFSVFAVPWFQYKILNKSC 112
DB 103 KTESNLVDCPFNEQPQCKERVICNFQINVAFLNKMSTNFC 145

RESULT 15
Q8K397 PRELIMINARY; PRT; 146 AA.
AC Q8K397;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE RIKEN cDNA 110017E11 gene (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027680; AAH27680.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
FT NON TER 1
SQ SEQUENCE 146 AA; 16380 MW; 9D77BB9A6063A5C4 CRC64;

Query Match 17.9%; Score 112.5; DB 11; Length 146;
Best Local Similarity 26.1%; Pred. No. 0.00027;
Matches 23; Conservative 20; Mismatches 38; Indels 7; Gaps 1;

QY 32 YNKESDDKYHFRIFRVLKVRQVTDHLEYHLNVEMQWTTCCQP-----ETTNCVQPQR 84
DB 56 YNMGSDSLYYFRDTKVYDAKYQLVAGIKYLYTLDIESTECKTRVSGEHMDLTTCPLAAG 115

QY 85 ELHKQVNCFFSVFAVPWFQYKILNKSC 112
DB 116 GQEKLRCNFELLEVFWNTTQLLKHC 143
```

RT "Functional annotation of a full-length mouse cDNA collection.;"

RL Nature 409:685-690(2001).  
 DR EMBL; AK020193; BAB32024.1; --  
 DR HSSP; P01034; IG96.  
 DR MGD; MGI:1925859; 8030411F24Rik.  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; Cy; 1.  
 SQ SEQUENCE 130 AA; 14947 MW; DD2F930B64B4E584 CRC64;

Query Match 19.2%; Score 121; DB 11; Length 130;  
 Best Local Similarity 27.5%; Pred. No. 2.9e-05;  
 Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 6 KTFLSVHEVMVENVYAKSLQWITDQYNKESDDKYHFRIVLKVQVQVTDHLEHVLNVE 65

DB 23 KEFLDVTKDL--DYFVASVEFAVAQFNDNNPEENTYKLLVEGQAQK-TWTMIFLMDLE 78

QY 66 MQWTTCCQKETT--NCVPERELHKQVNCFFSFAVPWFPEQYKILNKSC 112

DB 79 MGRITCKGHENIHNCPLQSGREKKVHCVFQVDARPFWSHFILSTC 127

RESULT 12

Q8VII3 ID Q8VII3 PRELIMINARY; PRT; 130 AA.

AC Q8VII3

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Cystatin SC.

GN 8030411F24Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=C57; TISSUE=Testis;

RA Li Y., Friel P.J., Griswold M.D.;

RT "Molecular cloning and characterization of cystatin SC and cystatin

TE-1, new members of the cystatin family.;"

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF440735; AAL30841.1; --

DR MGD; MGI:1925859; 8030411F24Rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; Cy; 1.

SQ SEQUENCE 130 AA; 15076 MW; DD34930B64AF5E8F CRC64;

Query Match 19.2%; Score 121; DB 11; Length 130;

Best Local Similarity 27.5%; Pred. No. 2.9e-05;

Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 6 KTFLSVHEVMVENVYAKSLQWITDQYNKESDDKYHFRIVLKVQVQVTDHLEHVLNVE 65

DB 23 KEFLDVTKDL--DYFVASVEFAVAQFNDNNPEENTYKLLVEGQAQK-TWTMIFLMDLE 78

QY 66 MQWTTCCQKETT--NCVPERELHKQVNCFFSFAVPWFPEQYKILNKSC 112

DB 79 MGRITCKGHENIHNCPLQSGREKKVHCVFQVDARPFWSHFILSTC 127

RESULT 13

Q9DAN8 ID Q9DAN8 PRELIMINARY; PRT; 128 AA.

AC Q9DAN8

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE 1700006F03Rik protein (Cystatin TE-1).

GN 1700006F03Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsumoto Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Stabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.;"  
 RL Nature 409:685-690(2001).  
 RN [2]  
 SEQUENCE FROM N.A.  
 RC STRAIN=C57; TISSUE=Testis;  
 RA Li Y., Friel P.J., Griswold M.D.;

RT "Molecular cloning and characterization of cystatin SC and cystatin

TE-1, new members of the cystatin family.;"

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK005670; BAB24179.1; --

DR EMBL; AF440737; AAL30843.1; --

DR MGD; MGI:1916612; 1700006F03Rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; Cy; 1.

SQ SEQUENCE 128 AA; 15036 MW; 40AFD00103E5E5D6 CRC64;

Query Match 18.2%; Score 114.5; DB 11; Length 128;

Best Local Similarity 29.7%; Pred. No. 0.00014;

Matches 27; Conservative 20; Mismatches 41; Indels 3; Gaps 2;

QY 24 SLQWITDQYNKESDDKYHFRIVLKVQVQVTDHLEHVLNVEQWTTCCQK--PETTNCVP 81

DB 36 SVEHVVFHFNENQDDDFAYKFLRVRSLRQKYT-LKYLVDLEMGRTLCGKYDEDIDNCPL 94

QY 82 QERELHKQVNCFFSFAVPWFPEQYKILNKSC 112

DB 95 QEGPGERKVRCTYIVETEAWTKFILNSTC 125

RESULT 14

Q9JM84 ID Q9JM84 PRELIMINARY; PRT; 148 AA.

AC Q9JM84

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE DD72 protein (Similar to cystatin 10) (Chondrocytes).

GN CST10 OR DD72

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

SEQUENCE FROM N.A.

```

RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bai J., Lao H., Ye X., Li Y., Lou J.;
RT "Molecular cloning and sequence analysis of cystatin cDNA from two
RL species of sturgeons.";
RL EMBL: AF334610; AAK16731.1; -.
DR HSSP: P01038; 1A90.
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART: SM00043; CY; 1.
DR PROSITE: PS00287; CYSTATIN; 1.
FT NON TER
SQ SEQUENCE 112 AA; 12231 MW; 48CEBFBED8A08C00 CRC64;

Query Match 22.1%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 2.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY 18 ENYAKDSLOWITDOYNKESDDKHFRIFRVLVKQVQVTDHLEHVLNVEWMTTCQK---P 74
DB 13 EGVQDALKEFAVEAFKASNDMTIHRVSKVKVQKQVAGIKIVTVQMGRTSCRGGAE 72

QY 75 ETTNC-----VPOBELHKQVNCFFSVFVAPVFPFQYKILNKSCS 113
DB 73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLKVNCT 112

RESULT 9
Q98SR3 ID Q98SR3 PRELIMINARY; PRT; 112 AA.
AC Q98SR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Acipenser schrenckii (Amur sturgeon).
OS Acipenseridae; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=111304;
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bai J., Lao H., Ye X., Li Y., Lou J.;
RT "Molecular cloning and sequence analysis of cystatin cDNA from two
RL species of sturgeons.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF334611; AAK16732.1; -.
DR HSSP: P01038; 1A90.
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR PROSITE: PS00287; CYSTATIN; 1.
FT NON TER
SQ SEQUENCE 112 AA; 12231 MW; 48CEBFBED8A08C00 CRC64;

Query Match 22.1%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 2.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY 18 ENYAKDSLOWITDOYNKESDDKHFRIFRVLVKQVQVTDHLEHVLNVEWMTTCQK---P 74
DB 13 EGVQDALKEFAVEAFKASNDMTIHRVSKVKVQKQVAGIKIVTVQMGRTSCRGGAE 72

QY 75 ETTNC-----VPOBELHKQVNCFFSVFVAPVFPFQYKILNKSCS 113
DB 73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLKVNCT 112

RESULT 10
Q98SR3 ID Q98SR3 PRELIMINARY; PRT; 112 AA.
AC Q98SR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Acipenser schrenckii (Amur sturgeon).
OS Acipenseridae; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=111304;
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bai J., Lao H., Ye X., Li Y., Lou J.;
RT "Molecular cloning and sequence analysis of cystatin cDNA from two
RL species of sturgeons.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF334611; AAK16732.1; -.
DR HSSP: P01038; 1A90.
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR PROSITE: PS00287; CYSTATIN; 1.
FT NON TER
SQ SEQUENCE 112 AA; 12231 MW; 48CEBFBED8A08C00 CRC64;

Query Match 22.1%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 2.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY 18 ENYAKDSLOWITDOYNKESDDKHFRIFRVLVKQVQVTDHLEHVLNVEWMTTCQK---P 74
DB 13 EGVQDALKEFAVEAFKASNDMTIHRVSKVKVQKQVAGIKIVTVQMGRTSCRGGAE 72

QY 75 ETTNC-----VPOBELHKQVNCFFSVFVAPVFPFQYKILNKSCS 113
DB 73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLKVNCT 112
```

```

Q8VIH8 ID Q8VIH8 PRELIMINARY; PRT; 130 AA.
AC Q8VIH8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin SC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RA Li Y., Friel P.J., Griswold M.D.;
RT "Molecular cloning and characterization of cystatin SC and cystatin
TE-1, new members of the cystatin family.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF442205; AAL35350.1; -.
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART: SM00043; CY; 1.
DR SEQUENCE 130 AA; 14981 MW; 7A752359860989C9 CRC64;

Query Match 19.4%; Score 122; DB 11; Length 130;
Best Local Similarity 27.5%; Pred. No. 2.3e-05;
Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 6 KTFLSVHEVMAVENYAKDSLOWITDOYNKESDDKHFRIFRVLVKQVQVTDHLEHVLNVE 65
DB 23 KEFLDVTKDQ---DYFVASVEFAVAFQNDNNSEENTYRLLEVGRAQKK-TWTMIFLMDLE 78

QY 66 MOWTTCQKPETT--NCVPOERELHKQVNCFFSVFVAPVFPFQYKILNKSC 112
DB 79 MERTICKKHDEHNHCPLLOGSGEKKVCHVCVQVDARPFWSHFVTLSTC 127

RESULT 11
Q9CX46 ID Q9CX46 PRELIMINARY; PRT; 130 AA.
AC Q9CX46;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 8030411F24RIK protein.
GN 8030411F24RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/60; TISSUE=Embryonic testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Niehi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
```

```
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048681; AA48681.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C/M.
DR Pfam; PF00031; cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 141 AA; 16825 MW; C20FA0D8A804951F CRC64;

Query Match      29.8%; Score 187.5; DB 11; Length 141;
Best Local Similarity 39.6%; Pred. No. 2.3e-12;
Matches 38; Conservative 20; Mismatches 35; Indels 3; Gaps 2;

QY 20 YAKDSLOWITDQNKESDDKYHFRIFVLKVRQVTDHLEYHLNVEMQWTTCC--PETT 77
Db 43 YVOHALWYAMKEYNKASNDLYNFRVVDILKSQSDITDSLEYLYLEVNARTMCKKIADGNE 102

QY 78 NCVPQERELHKQVNCFFSVFVFPWFQYKILNKSC 112
Db 103 NCLFQDDPKMKVFCIFIVSSKFWKELKWLKQC 138

RESULT 6
Q9EPX9 PRELIMINARY; PRT; 140 AA.
AC Q9EPX9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=21010502; PubMed=11144350;
RA Taupin P.J., Ray J., Fischer W.H., Suhr S.T., Hakansson K., Grubb A.,
RA Gage P.H.;
RA "FGF-2-Responsive neural stem cell proliferation requires Ccg, a novel
RT autocrine/paracrine cofactor.";
RL Neuron 28:385-397(2000).
DR HSPB; AF311741; AAC40283.1; -.
DR HSPB; P01034; I996.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT CHAIN 21 140 CYSTATIN C.
FT VARIANT 16 16 A -> G.
FT VARIANT 84 84 L -> F.
SQ SEQUENCE 140 AA; 15517 MW; 3A563406DD58D795 CRC64;

Query Match      26.5%; Score 166.5; DB 11; Length 140;
Best Local Similarity 33.3%; Pred. No. 4.1e-10;
Matches 35; Conservative 24; Mismatches 43; Indels 3; Gaps 2;

QY 13 EVMAVENYAKDSLOWITDQNKESDDKYHFRIFVLKVRQVTDHLEYHLNVEMQWTTCC 72
Db 35 EADANEGVRALDFAVSEYKNGSNDAYHGRATQVRRARQLVAGVNYFLDVEGRTTCT 94

QY 73 KPET--TNC-VPOERELHKQVNCFFSVFVFPWFQYKILNKSCSS 114
Db 95 KSQTNLTDCFFHDQPHLMRRALCSFQIYSVPWKGTSLTNFSCN 139

RESULT 7
Q80Y72 PRELIMINARY; PRT; 140 AA.
ID Q80Y72
Q80Y72
```

```
AC Q80Y72;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cystatin-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=22388237; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Scheetz T.E.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullaby S.J.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048646; AA48646.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C/M.
DR Pfam; PF00031; cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 140 AA; 16199 MW; 32633E99C4697DA0 CRC64;

Query Match      23.8%; Score 149.5; DB 11; Length 140;
Best Local Similarity 30.4%; Pred. No. 2.7e-08;
Matches 31; Conservative 27; Mismatches 41; Indels 3; Gaps 2;

QY 13 EVMAVENYAKDSLOWITDQNKESDDKYHFRIFVLKVRQVTDHLEYHLNVEMQWTTCC 72
Db 37 KAMSKKN-INSTLHFFIRSYNNASNDTYLYQVKLIQQGQMLTGVLYLVTVIGRTCK 95

QY 73 KPET--TNCVPQERELHKQVNCFFSVFVFPWFQYKILNKSC 112
Db 96 KNETKACSPLOSSKLUKSLICKSLIYSVPMWNYQLWNNSC 137

RESULT 8
Q98SR4 PRELIMINARY; PRT; 112 AA.
ID Q98SR4
AC Q98SR4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin (Fragment).
OS Acipenser sinensis (Chinese sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=61970;
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 68.9087 Seconds  
(without alignments)  
526.560 Million cell updates/sec

Title: US-09-941-314-3

Perfect score: 629

Sequence: 1 YQARKKTFSLVHVMVAVENY.....VFAPWFEQYKILNKSCSSD 115

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriopl:\*

17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	100.0	138	4 Q8WXU6	Q8wxu6 homo sapien
2	400.5	63.7	103	4 Q8WXU5	Q8wxu5 homo sapien
3	360	57.2	139	11 Q8K5A3	Q8k5a3 rattus norv
4	187.5	29.8	141	11 Q9DAP1	Q9dap1 mus musculu
5	187.5	29.8	141	11 Q80ZN5	Q80zn5 mus musculu
6	166.5	26.5	140	11 Q8EPX9	Q8epx9 mus musculu
7	149.5	23.8	140	11 Q80Y72	Q80y72 mus musculu
8	139	22.1	112	13 Q98SR4	Q98sr4 acipenser s
9	139	22.1	112	13 Q98SR3	Q98sr3 acipenser s
10	122	19.4	130	11 Q8VH8	Q8vhi8 rattus norv
11	121	19.2	130	11 Q9CX46	Q9cx46 mus musculu
12	121	19.2	130	11 Q8VH13	Q8vhi3 mus musculu
13	114.5	18.2	128	11 Q9DAN8	Q9dan8 mus musculu
14	113	18.0	148	11 Q9JMH4	Q9jmh4 mus musculu
15	112.5	17.9	146	11 Q8K397	Q8k397 mus musculu
16	112.5	17.9	149	11 Q8VHC1	Q8vhc1 rattus norv

Q9dib1 mus musculu  
Q7z4j8 homo sapien  
Q8vii2 rattus norv  
Q8jfu5 brachydanio  
Q9ql5 mus musculu  
Q9ty65 onchocerca  
O16159 brugia mala  
Q9nh95 litomosoida  
Q9d264 mus musculu  
Q25620 onchocerca  
P90698 brugia mala  
Q12700 debaryomyce  
Q17108 acanthochei  
Q9u9a1 onchocerca  
Q8mvb6 ixodes scap  
O16454 caenorhabdi  
Q8ibf4 plasmodium  
Q8b381 rattus norv  
Q8ndm7 homo sapien  
Q95k43 macaca fasc  
Q8i5x0 plasmodium  
Q7ytp6 sus scrofa  
Q8wxk6 homo sapien  
P70517 rattus norv  
Q8wxx7 homo sapien  
Q8ib61 plasmodium  
Q8i160 bacillus ce  
Q49557 mycoplasma  
Q9ur18 candida alb

#### ALIGNMENTS

RESULT 1

Q8WXU6 PRELIMINARY; PRT; 138 AA.

AC Q8WXU6; DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE SC13.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

ON NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;

RT "SC13: A novel epididymal specific member of the cystatin family.";

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF335480; AAL7191.1; -.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; Cy; 1.

SQ SEQUENCE 138 AA; 16506 MW; E49440ACA3585C64 CRC64;

Query Match 100.0%; Score 629; DB 4; Length 138;

Best Local Similarity 100.0%; Pred. No. 8.9e-60;

Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YQARKKTFSLVHVMVAVENYAKDSLOWITDQYNKSDDKYHFRIVLKVQVQVTDHLEY 60

Db 24 YQARKKTFSLVHVMVAVENYAKDSLOWITDQYNKSDDKYHFRIVLKVQVQVTDHLEY 83

Qy 61 HLNVEMQWTTCCQPETTNCVQRELHKQVNCFFSVFVFWFQYKILNKSCSSD 115

Db 84 HLNVEMQWTTCCQPETTNCVQRELHKQVNCFFSVFVFWFQYKILNKSCSSD 138

RESULT 2

Q8WXU5



```

AC 019092;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN CST3.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA "Cystatin C. Icelandic-like mutation in an animal model of
RT cerebrovascular beta-amyloidosis."
RL Stroke 27:2080-2085(1996).
CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U51912; BAB64050.1; -.
DR HSP; P01034; I396.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26
FT CHAIN 27 146 CYSTATIN C.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 15857 MW; F0B3BB774A29DF26 CRC64;

Query Match 24.6%; Score 155; DB 1; Length 146;
Best Local Similarity 33.7%; Pred. No. 5.1e-09;
Matches 34; Conservative 21; Mismatches 42; Indels 4; Gaps 3;

QY 18 ENYAKDSLOWITDQYNKESDDKTHFRIFRVLVKQVQVTDHLEYHLNVEMQWTTQK--PE 75
Db |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
46 EGVRRALDPVAVSEYKNASNDYHSRALQVVRARQIVAGVNYFLDVELGRITCTKTQPN 105
QY 76 TTNCVQER-ELHKQVNCFFSVAPWPFQYKILNKCSSD 115
Db |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
106 LDNCPFHEQPHLRKAKFCFQIYVTPW-QGTMTLSKSTCD 145

```

Search completed: March 18, 2004, 14:16:08  
Job time : 14.6905 secs



RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehesvaikho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConnachie L.J., McIlroy K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Suze C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
CC -I- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -I- SIMILARITY: Belongs to the cystatin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; AL096677; CAC03440.2; --  
DR HSPSP; P01038; I490.  
DR Gnew; HGNC:15958; CST1L.  
DR InterPro; IPR000010; Cystatin.  
DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
KW Thiol protease inhibitor; Signal.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 165 CYSTATIN-LIKE 1.  
FT SITE 93 97 SECONDARY AREA OF CONTACT (POTENTIAL).  
FT DISULFID 111 121 BY SIMILARITY.  
FT DISULFID 134 154  
FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 54 54 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 165 AA; 19312 MW; 9D66D685875DAEEA CRC64;

Query Match 25.0%; Score 157.5; DB 1; Length 165;  
Best Local Similarity 26.2%; Pred. No. 3.2e-09;  
Matches 32; Conservative 30; Mismatches 37; Indels 23; Gaps 3;  
QY 13 EVMAVENYAKSLQWITDQNKESDDKYHFRIFVLKVQ-----R 52  
Db 34 KLSMKKN-MNSTLNFFQSTNNASNDLYLRVORLKRKSQERVSHVLMGLVHNTSTDSR 92  
QY 53 QVTDHLYHLNVEMQWTTCKPPTN--CVFQREHLKQVNCFFSFAVPFWFQYKILNK 110  
Db 93 QLTTGVEYIVTKIGWTKCKRNDTSNSSLQSKLRKSLICESLIYTMPVNFQVQLWN 152  
QY 111 SC 112  
Db 153 SC 154

RESULT 14  
CYTC\_HUMAN  
ID CYTC\_HUMAN STANDARD; PRT; 146 AA.  
AC P01034;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)  
DE (Post-gamma-globulin).  
DE CST3.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=87219149; PubMed=3495457;  
RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;  
RT "Molecular cloning and sequence analysis of cDNA coding for the  
RT precursor of the human cysteine proteinase inhibitor cystatin C.";  
RL FEBS Lett. 216:229-233 (1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Leukocyte;  
RX MEDLINE=90303202; PubMed=2363674;  
RA Abrahamson M., Olafsson I., Palsdottir A., Ulvsbaeck M., Lundwall A.,  
RA Jansson O., Grubb A.;  
RT "Structure and expression of the human cystatin C gene.";  
RL Biochem. J. 268:287-294 (1990).  
RN [3]  
RP SEQUENCE FROM N.A. (HCHWA VARIANT).  
RC TISSUE=Brain;  
RX MEDLINE=89235594; PubMed=2541223;  
RA Levy E., Lopez-Otin C., Ghiso J., Geltner D., Frangione B.;  
RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is  
RT related to a mutation in the cystatin C gene, an inhibitor of  
RT cysteine proteases.";  
RL J. Exp. Med. 169:1771-1778 (1989).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89350949; PubMed=2764935;  
RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,  
RA Isomura S., Sanada K.;  
RT "The human cystatin C gene (CST3) is a member of the cystatin gene  
RT family which is localized on chromosome 20.";  
RL Biochem. Biophys. Res. Commun. 162:1324-1331 (1989).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX Dickinson D.P., Hewett-Emmett D., Thiesse M.;  
RT "Acquisition of complex patterns of differential expression in  
RT epithelial cell populations during the evolution of type 2 cystatin  
RT genes.";  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehesvaikho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConnachie L.J., McIlroy K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,



Query Match	25.5%	Score 160.5;	DB 1;	Length 140;
Best Local Similarity	32.4%;	Pred. No. 1.3e-09;		
Matches 34;	Conservative 24;	Mismatches 44;	Indels 3;	Gaps 2;



```

OX NCBI_TaxID=10116;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=Buffalo;
RX MEDLINE=90092122; PubMed=2689174;
RA Cole T., Dickson P.W., Ennard F., Averill F., Risbridger G.,
RT "The cDNA structure and expression analysis of the genes for the
RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
RT in rat brain.";
RL Eur. J. Biochem. 186:35-42(1989).
RN [2]
RP SEQUENCE OF 8-127.
RX MEDLINE=90380276; PubMed=2400577;
RA Ennard F., Ennard A., Faucher D., Capony J.-P., Derancourt J.,
RA Brillard M., Gauthier F.;
RT "Rat cystatin C: the complete amino acid sequence reveals a site for
RT N-glycosylation.";
RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).
RN [3]
RP SEQUENCE OF 8-49.
RX MEDLINE=88313020; PubMed=3044831;
RA Ennard A., Ennard F., Faucher D., Gauthier F.;
RT "Two rat homologues of human cystatin C.";
RL FEBS Lett. 236:475-478(1988).
RN [4]
RP SEQUENCE OF 8-20.
RC TISSUE=Sertoli cells;
RX MEDLINE=92225121; PubMed=1563513;
RA Ennard A., Ennard F., Guillou F., Gauthier F.;
RT "Production of the cysteine proteinase inhibitor cystatin C by rat
RT Sertoli cells.";
RL FEBS Lett. 300:131-135(1992).
CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity. Known to inhibits cathepsin B,
CC H, and L.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16957; CAA34831.1; -.
DR PIR; S07085; S07085.
DR PIR; S10587; S10587.
DR HSSP; P01034; IG96.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; Cy; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Signal.
FT NON_TER 1 1
FT SIGNAL <1 7
FT CHAIN 8 127 CYSTATIN C.
FT ACT_SITE 18 18 REACTIVE SITE.
FT SITE 62 66 SECONDARY AREA OF CONTACT.
FT DISULFID 80 90 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT CONFLICT 25 25 A -> E (IN REF. 2).
SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;
Query Match 27.3%; Score 171.5; DB 1; Length 127;
Best Local Similarity 33.3%; Pred. No. 8.9e-11;
Matches 35; Conservative 24; Mismatches 43; Indels 3; Gaps 2;
QY 13 EVMAVENYAKDSLQWTDQYNKESDDKYPFRFLVKVQRQVTDHLEYHLNVMQWTTQC 72
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
22 EADASEGVQRALDFAVSEYKNGSDAYHSRAIQVVRARKQLVAGINYLDVEMGRITCT 81

```

```

QY 73 KBET--TNC-VPOERELHKVNCFFSVFAVPPEQYKILNKSCSS 114
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
82 KSGTNLTNCFHDPHLMKALCSFOIYSVPWKGTHTLTKSSCKN 126
RESULT 8
CYT_CHK CK
ID_CYT_CHK STANDARD; PRT; 139 AA.
AC P01038;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin precursor (Egg-white cystatin).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Gallus; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90008873; PubMed=2793849;
RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;
RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,
RT and tissue distribution.";
RL J. Biol. Chem. 264:17164-17169(1989).
RN [2]
RP SEQUENCE OF 24-139.
RX MEDLINE=84178305; PubMed=6712597;
RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;
RT "Cystatin. Amino acid sequence and possible secondary structure.";
RL Biochem. J. 217:813-817(1984).
RN [3]
RP SEQUENCE OF 24-139.
RX MEDLINE=84110059; PubMed=6662498;
RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,
RA Machleidt W.;
RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
RT of cystatin from chicken egg white.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
RN [4]
RP CHARACTERIZATION OF PROTEIN.
RX MEDLINE=83256421; PubMed=6409085;
RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,
RA Sunter D.C., Barrett A.J.;
RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved
RT purification from egg white, characterization, and detection in
RT chicken serum.";
RL Biochem. J. 211:129-138(1983).
RN [5]
RP DISULFIDE BONDS.
RA Grubb A., Loeffberg H., Barrett A.J.;
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
RT cystatin.";
RL FEBS Lett. 170:370-374(1984).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=89252033; PubMed=2721673;
RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,
RA Bode W.;
RT "The cysteine proteinase inhibitor chicken cystatin is a
RT phosphoprotein.";
RL FEBS Lett. 248:162-168(1989).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=89052676; PubMed=3191914;
RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshikov A.,
RA Brzin J., Kos J., Turk V.;
RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and
RT its possible mode of interaction with cysteine proteinases.";
RL EMBO J. 7:2593-2599(1988).
RN [8]
RP STRUCTURE BY NMR.

```

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Frazer A., French L., Garner P.,  
 RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.B., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leheslahti M.H., Leversha M.A., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Suleston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation. LOCATION: Secreted.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epithelium.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF059244; AAC14707.1; -;  
 DR EMBL; AL109954; CAB64234.1; -;  
 DR HSSP; P01034; IG96.  
 DR Genew; HGNC:2480; CST8.  
 DR GO; GO:0004869; F:Cysteine protease inhibitor activity; TAS.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR Thiol protease inhibitor; Signal; Polymorphism.  
 FT SIGNAL 1 21  
 FT CHAIN 22 142  
 FT SITE 77 81  
 FT SITE 95 105  
 FT DISULFID 119 139  
 FT DISULFID 27 27  
 FT CARBOHYD 39 39  
 FT VARIANT 142 142  
 FT SEQUENCE 142 AA; 16275 MW; 9A3512757E0F4ECD CRC64;  
 Query Match 30.0%; Score 188.5; DB 1; Length 142;  
 Best Local Similarity 40.4%; Pred. No. 1.8e-12;  
 Matches 38; Conservative 23; Mismatches 30; Indels 3; Gaps 2;  
 QY 22 KDSLOWITDOYNKESDDKYHFRVLKQVROVTDHLEHYHLNVMQWTTCKPETTNC-C 79  
 DB 46 KQCLWFAMQYKNKESDDKYHFRVLKQVROVTDHLEHYHLNVMQWTTCKPETTNC 105  
 QY 80 VPQER-ELHKQVNCFFSFVAFVFPFQYKILNKSC 112  
 DB 106 AIQENSKLKKRLKSCSFVLGALPWFNGEFTVMEKKC 139

RESULT 6  
 CYTC\_RABIT  
 ID CYTC\_RABIT STANDARD; PRT; 148 AA.  
 AC O97862;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin C precursor.  
 GN CST3.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OC NCBI\_TaxID=9986;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN=Japanese white; TISSUE=Bone;  
 RX MEDLINE=98424349; PubMed=9753427;  
 RA Kobori M., Ikeda Y., Nara H., Kato M., Kumegawa M., Nojima H.,  
 RA Kawashima H.;  
 RT "Large scale isolation of osteoclast-specific genes by an improved  
 RT method involving the preparation of a subtracted cDNA library.";  
 RL Genes Cells 3:459-475(1998).  
 CC -1- FUNCTION: This is a thiol proteinase inhibitor.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB009342; BAA75921.1; -;  
 DR HSSP; P01034; IG96.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYPSTATIN; FALSE\_NEG.  
 DR Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 28  
 FT CHAIN 29 148  
 FT SITE 39 39  
 FT SITE 83 87  
 FT DISULFID 101 111  
 FT DISULFID 125 145  
 FT SEQUENCE 148 AA; 16346 MW; 1523C8311695B9A CRC64;  
 Query Match 27.7%; Score 174.5; DB 1; Length 148;  
 Best Local Similarity 30.8%; Pred. No. 5.2e-11;  
 Matches 33; Conservative 30; Mismatches 41; Indels 3; Gaps 2;  
 QY 11 VHEVMAVENYAKDSLOWITDOYNKESDDKYHFRVLKQVROVTDHLEHYHLNVMQWTT 70  
 DB 41 LEBVDAQKDVQKALGFASEYKNGSDRYHRSALQVVRARRQLVSGVKYLDVIGRTT 100  
 QY 71 CQKPEP--TNC-VPOREHLKQVNCFFSFVAFVFPFQYKILNKSCSS 114  
 DB 101 CTKTQTNLANCPHDPDQLQKMLCSFEIYVPLNKLISLKSQCQN 147

RESULT 7  
 CYTC\_RAT  
 ID CYTC\_RAT STANDARD; PRT; 127 AA.  
 AC P14841;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin C precursor (fragment).  
 GN CST3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```
CC maturation.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
CC expression in the testis. Within the testis it is localized to the
CC elongating spermatids, whereas within the epididymis it is
CC exclusively synthesized by the proximal caput epithelium.
CC -!- INDUCTION: Testicular factors or hormones other than androgens
CC present in the testicular fluid may be involved in the regulation
CC of CRE gene expression
CC -!- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF091503; AAC61754.1; -.
DR EMBL; AF090691; AAC36316.1; -.
DR EMBL; S49926; AAC35390.1; -.
DR PIR; A45361; A45361.
DR HSSP; P01034; IG96.
DR MGD; MGI:107161; Gsta8.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; Cy; 1.
DR Thiol protease inhibitor; Signal.
KW SIGNAL 1 19
FT CHAIN 20 142
FT SITE 77 81 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
FT DISULFID 95 105 PROTEIN.
FT DISULFID 119 139 SECONDARY AREA OF CONTACT (POTENTIAL).
FT CARBOHYD 39 39 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 4 15 PLWSLILFIIP -> GTRDEOVGESQK (IN REF. 2).
SQ SEQUENCE 142 AA; 16288 MW; 50B446E98F66673E CRC64;

Query Match 34.0%; Score 214; DB 1; Length 142;
Best Local Similarity 37.1%; Pred. No. 4.5e-15;
Matches 43; Conservative 29; Mismatches 30; Indels 14; Gaps 4;

QY 10 SVHEWAVENY-----AKDSLOWITDQYNKESDDKYHFRIFRVLVKQVQVTDHLE 59
DB 25 SKNEVKA-QNYFGISINISNANVKQVWFAMKEYNKESEDKYFLVDKILHAKLQITDRME 83
QY 60 YHLNVEMQWTTCKP--ETTNCVPOER-ELHKQVNCFFSVFVFPWFQYKILNKSC 112
DB 84 YQIDVQISRNCCKPLNNTENCIPQKPKPELEKXKSCFLVGLPWNGEFLLSKEC 139

RESULT 4
CST8_RAT
ID CST8_RAT STANDARD; PRT; 142 AA.
AC O88969;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
DE 8).
GN CST8 OR CRES.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Epididymis;
RC MEDLINE=99247899; PubMed=10229662;
RA Cornwall G.A., Hsiao N., Sutton H.G.;
RA "Structure, alternative splicing and chromosomal localization of the
```

```
RT cystatin-related epididymal spermatogenic gene.";
RL Biochem. J. 340:85-93(1999).
CC -!- FUNCTION: Performs a specialized role during sperm development and
CC maturation.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF090692; AAC36317.1; -.
DR HSSP; P01034; IG96.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; Cy; 1.
DR Thiol protease inhibitor; Signal.
KW SIGNAL 1 19
FT CHAIN 20 142
FT SITE 77 81 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
FT DISULFID 95 105 PROTEIN.
FT DISULFID 119 139 SECONDARY AREA OF CONTACT (POTENTIAL).
FT CARBOHYD 100 100 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 142 AA; 16246 MW; FB873FAA6B6CAE34 CRC64;

Query Match 31.6%; Score 199; DB 1; Length 142;
Best Local Similarity 35.7%; Pred. No. 1.5e-13;
Matches 41; Conservative 26; Mismatches 36; Indels 12; Gaps 4;

QY 10 SVHEWAVENYAKDSLQ-----WIT-DQYNKESDDKYHFRIFRVLVKQVQVTDHLEY 60
DB 25 SKNEVKAQRYFGSISISNANVKQVWFAMKEYNKGSEDKYFLDLTKLHATLQITDRMEY 84
QY 61 HLNVEWQWTTCKP--ETTNCVPOER-ELHKQVNCFFSVFVFPWFQYKILNKSC 112
DB 85 HIDVQISRNCCKPLNNTENCIPQKPKPELEKXKSCFLVGLPWNGEFLLSKEC 139

RESULT 5
CST8_HUMAN
ID CST8_HUMAN STANDARD; PRT; 142 AA.
AC O60676;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
DE 8).
GN CST8 OR CRES.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95344753; PubMed=7619504;
RA Cornwall G.A., Hann S.R.;
RT "Transient appearance of CRES protein during spermatogenesis and
RT caput epididymal sperm maturation.";
RL Mol. Reprod. Dev. 41:37-46(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AL096677; CAC13170.1; -.  
 HSSP; AL096677; CAC17423.1; -.  
 HSSP; P01038; 1A90.  
 Genew; HGNC:15959; Cst11.  
 InterPro; IPR000010; Cystatin.  
 Pfam; PF00031; Cystatin; 1.  
 SMART; SM00043; Cy; 1.  
 PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
 Thiol protease inhibitor; Signal; Alternative splicing.  
 SIGNAL 1 25 POTENTIAL.  
 CHAIN 26 137 SECONDARY AREA OF CONTACT (POTENTIAL).  
 SITE 75 79 BY SIMILARITY.  
 DISULFID 93 101 BY SIMILARITY.  
 DISULFID 114 134 BY SIMILARITY.  
 CARBOHYD 131 131 N-LINKED (GLCNAC... ) (POTENTIAL).  
 VARSPLIC 76 110 Missing (in isoform 2).  
 FTID-VSP 001260.  
 SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 629; DB 1; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-57;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKKTFSLVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRFVLKQVQVTDHLEY 60  
 DB 23 YQARKKTFSLVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRFVLKQVQVTDHLEY 82

QY 61 HLNVENQWTTCKPPTTNCVQPERLHKQVNCFFSVFVFPWFQYKILNKSCSSD 115  
 DB 83 HLNVENQWTTCKPPTTNCVQPERLHKQVNCFFSVFVFPWFQYKILNKSCSSD 137

Query Match 59.1%; Score 372; DB 1; Length 139;  
 Best Local Similarity 54.9%; Pred. No. 3e-31;  
 Matches 62; Conservative 29; Mismatches 22; Indels 0; Gaps 0;

QY 1 YQARKKTFSLVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRFVLKQVQVTDHLEY 60  
 DB 24 YQVQKRTFIRIEVSALESSVETLEYVDYVDEYNKXSDLYNFRILKIMKQVTHLEY 83

QY 61 HLNVENQWTTCKPPTTNCVQPERLHKQVNCFFSVFVFPWFQYKILNKSCS 113  
 DB 84 HITVEMQRTTCLKTSTSLCDIQKGLHKKIQCYFSVYAIPTWVEVFKILKNCT 136

RESULT 3  
 ID CS11\_MOUSE STANDARD; PRT; 139 AA.  
 AC Q9D269;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin 11 precursor.  
 GN Cst11.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bareh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtauki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."

QY 1 YQARKKTFSLVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRFVLKQVQVTDHLEY 60  
 DB 24 YQVQKRTFIRIEVSALESSVETLEYVDYVDEYNKXSDLYNFRILKIMKQVTHLEY 83

QY 61 HLNVENQWTTCKPPTTNCVQPERLHKQVNCFFSVFVFPWFQYKILNKSCS 113  
 DB 84 HITVEMQRTTCLKTSTSLCDIQKGLHKKIQCYFSVYAIPTWVEVFKILKNCT 136

RESULT 3  
 ID CS11\_MOUSE STANDARD; PRT; 142 AA.  
 AC P32766; O89102;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-related epididymal specific protein) (Cystatin 8).  
 GN CST8 OR CRES.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H, and CD-1;  
 RX MEDLINE=99247899; PubMed=10229662;  
 RA Cornwall G.A., Hsiao N., Sutton H.G.;  
 RT "Structure, alternative splicing and chromosomal localization of the cystatin-related epididymal spermatogenic gene."  
 RL Biochem. J. 340:85-93 (1999).  
 RN [2]  
 RP SEQUENCE OF 4-142 FROM N.A.  
 RC TISSUE=Epididymis;  
 RX MEDLINE=93078799; PubMed=1280328;  
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;  
 RT "The CRES gene: a unique testis-regulated gene related to the cystatin family is highly restricted in its expression to the proximal region of the mouse epididymis."  
 RL Mol. Endocrinol. 6:1653-1664 (1992).  
 CC -!- FUNCTION: Performs a specialized role during sperm development and

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 13.6905 Seconds  
(without alignments)  
437.389 Million cell updates/sec

Title: US-09-941-314-3  
Perfect score: 629  
Sequence: 1 YQARKKTFSLVHEVMAVENVY.....VFAVPWFQYKLNKSCSD 115

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	629	100.0	137	CS11 HUMAN	Q9h112 homo sapien
2	372	59.1	139	CS11 MOUSE	Q9d269 mus musculus
3	214	34.0	142	CST8 MOUSE	P32766 mus musculus
4	199	31.6	142	CST8 RAT	O88969 rattus norv
5	188.5	30.0	142	CST8 HUMAN	O60676 homo sapien
6	174.5	27.7	148	CYTC RABIT	O97862 oryctolagus
7	171.5	27.3	127	CYTC RAT	P14841 rattus norv
8	170.5	27.1	139	CYT CHICK	P01038 gallus gall
9	169.5	26.9	116	CYT CONJA	P81061 coturnix co
10	160.5	25.5	140	CYTC MOUSE	P21460 mus musculus
11	160.5	25.5	141	CYTC HUMAN	P09228 homo sapien
12	160	25.4	146	CYTC SAISC	O19093 salmifir sci
13	157.5	25.0	165	CSTL HUMAN	Q9h114 homo sapien
14	155	24.6	146	CSTL HUMAN	P01034 homo sapien
15	155	24.6	146	CYTC MACMU	O19092 macaca mula
16	153.5	24.4	148	CYTC BOVIN	P01035 bos taurus
17	146.5	23.3	141	CYTC HUMAN	P01036 homo sapien
18	143	22.7	129	CYT CYPCA	P35481 cyprinus ca
19	139.5	22.2	141	CYTN HUMAN	P01037 homo sapien
20	135	21.5	147	CST9 HUMAN	Q9h491 homo sapien
21	130.5	20.7	142	CYTD HUMAN	P28325 homo sapien
22	129.5	20.6	141	CYTS RAT	P19313 rattus norv
23	127	20.2	111	CYT BITAR	P08935 bitis ariet
24	127	20.2	130	CYT ONCMY	Q91195 oncorhynch
25	127	20.2	137	CST9 MOUSE	Q9z0h6 mus musculus
26	126	20.0	130	CYT ONCKE	Q98967 oncorhynch
27	121.5	19.3	149	CYTM HUMAN	O15828 homo sapien
28	109	17.3	145	CYTF HUMAN	O76096 homo sapien
29	98.5	15.7	434	KNL2 BOVIN	P01047 bos taurus
30	98.5	15.7	619	KNH2 BOVIN	P01045 bos taurus
31	98	15.6	144	CYTF MOUSE	O89098 mus musculus
32	96.5	15.3	436	KNL1 BOVIN	P01046 bos taurus
33	96.5	15.3	621	KNH1 BOVIN	P01044 bos taurus

34	95.5	15.2	162	1	CYTX ONCVO	P22085 onchocerca
35	93.5	14.9	644	1	KNG HUMAN	P01042 homo sapien
36	87	13.8	639	1	KNG RAT	P08934 rattus norv
37	82.5	13.1	474	1	SYE_VIBPA	Q87r16 vibrio para
38	81	12.9	430	1	KNT2 RAT	P08932 rattus norv
39	79	12.6	430	1	KNT1 RAT	P01048 rattus norv
40	79	12.6	661	1	KNG MOUSE	O08677 mus musculus
41	77	12.2	214	1	CAPA_KLUULA	O74232 kluyveromyc
42	77	12.2	430	1	MESJ_HAEIN	P44689 haemophilus
43	75.5	12.0	474	1	SYE_VIBVY	Q7mmw8 vibrio vuln
44	75	11.9	400	1	BGIB HUMAN	Q06430 homo sapien
45	73.5	11.7	438	1	EXG_CANAL	P29717 candida alb

ALIGNMENTS

```
RESULT 1
CS11_HUMAN STANDARD; PRT; 137 AA.
AC O9H112; O9H113;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Cystatin 11 precursor.
GN CST11 OR CST8L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knighs A., Laird G.K., Lawlor S.,
RA Lehaeslaio M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O9H112-1; SequencesDisplayed;
CC Name=2;
CC IsoId=O9H112-2; SequencesVSP 001260;
CC Note=No experimental confirmation available;
CC -1- SIMILARITY: Belongs to the cystatin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```





C;Comment: The intracellular role of this protein is the inhibition of intralysosomal p

Query Match 21.3%; Score 134; DB 1; Length 111;  
Best Local Similarity 27.4%; Pred. No. 2.1e-06;  
Matches 26; Conservative 27; Mismatches 40; Indels 2; Gaps 1;

Qy 22 KDSLOWITDOYNKESDDKYHFRIFRVLKQVQYTDHLEHYHLNVEMQWTTCK--PETTNC 79  
Db 17 RDLAQFAVHEHNKKTDMFVRQAKVNAQKVQVSGMKYIFTVMGRTPCRKGGVERICS 76  
Qy 80 VPQERELHKQVNCFFSVFAVPWFPEQYKILNKSCSS 114  
Db 77 VHKDPQMAVPYKCTFEWWSIPMWSIGIKWKNQCES 111

RESULT 12  
A47142  
Cystatin D precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 16-Jul-1999  
C;Accession: A47142; S18212  
R;Freije, J.P.; Balbin, M.; Abrahamson, M.; Velasco, G.; Dalboge, H.; Grubb, A.; Lopez-C  
J. Biol. Chem. 268, 15737-15744, 1993  
A;Title: Human cystatin D. cDNA cloning, characterization of the Escherichia coli expres  
A;Reference number: A47142; MUID:93340179; PMID:8340398  
A;Accession: A47142  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-142 <PRE>  
A;Cross-references: GB:W0377; NID:G398710; PIDN:CAA49838.1; PID:G398711  
A;Note: Single residue difference between this report and S18218 was investigated and sh  
R;Freije, J.P.; Abrahamson, M.; Olafsson, I.; Velasco, G.; Grubb, A.; Lopez-Otin, C.  
J. Biol. Chem. 266, 20538-20543, 1991  
A;Title: Structure and expression of the gene encoding cystatin D, a novel human cystein  
A;Reference number: S18212; MUID:92041895; PMID:1939105  
A;Accession: S18212  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-45; 'C', 47-142 <PR2>  
A;Cross-references: EMBL:X59964; NID:G30263; PIDN:CAA42590.1; PID:G30264  
C;Genetics:  
A;Gene: GDB:CS75  
A;Cross-references: GDB:136380; OMIM:123858  
A;Map position: 20p11.21-20p11.21  
A;Introns: 77/3; 115/3  
C;Superfamily: cystatin; cystatin homology  
C;Keywords: cysteine proteinase inhibitor; extracellular protein; saliva  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;30-142/Domain: cystatin homology <CYS>

Query Match 20.9%; Score 131.5; DB 2; Length 142;  
Best Local Similarity 27.8%; Pred. No. 5e-06;  
Matches 32; Conservative 25; Mismatches 53; Indels 5; Gaps 4;

Qy 3 ARKKTPL-SVHEWAVENYAKDSLQWITDOYNKE-SDDKYHFRIFRVLKQVQYTDHLEY 60  
Db 25 AQSRITAGGIHATDLNDSQVRLDFAISEYNKVINKDEYYSRPLQWMAAYQQIVGVNY 84  
Qy 61 HLANVEMQWTTCK--PETTNC-VPQERELHKQVNCPSVFAVPWFPEQYKILNKSC 112  
Db 85 YFNVKVGRGTRTCTKSQPNLDNCPNDPQKLKEEFCSQINEVPWEDKISILNYKC 139

RESULT 13  
JQ1470  
Cystatin S precursor - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 16-Jul-1999

C;Accession: JQ1470; A31891; S05252; S15132  
R;Cox, J.L.; Shaw, P.A.  
Gene 110, 175-180, 1992  
A;Title: Structure, organization and regulation of a rat cysteine proteinase inhibitor-er  
A;Reference number: JQ1470; MUID:92165056; PMID:1537554  
A;Accession: JQ1470  
A;Molecule type: DNA  
A;Residues: 1-141 <COX>  
A;Cross-references: GB:M75281; NID:G294537; PIDN:AAA41068.1; PID:G294538  
R;Shaw, P.A.; Cox, J.L.; Barka, T.; Naito, Y.  
J. Biol. Chem. 263, 18133-18137, 1988  
A;Title: Cloning and sequencing of cDNA encoding a rat salivary cysteine proteinase inhibit  
A;Reference number: A31891; MUID:89053983; PMID:3263967  
A;Accession: A31891  
A;Molecule type: mRNA  
A;Residues: 'ISTDYLXISFHEHETLSC', 27-141 <SHA>  
R;Bedi, G.S.  
Arch. Biochem. Biophys. 273, 245-253, 1989  
A;Title: Amino acid sequence of an inducible cysteine proteinase inhibitor (cystatin) fr  
A;Reference number: S05252; MUID:89334379; PMID:2757396  
A;Accession: S05252  
A;Molecule type: protein  
A;Residues: 28-113, 'QE', 116-141 <BED>  
A;Note: 87-Glu and 88-His were also found; disulfide bonds determined  
R;Nishiura, T.; Ishibashi, K.; Abe, K.  
Biochim. Biophys. Acta 1077, 346-354, 1991  
A;Title: Isolation of three forms of cystatin from submandibular saliva of isoproterenol-  
A;Reference number: S15132; MUID:91230145; PMID:1903068  
A;Accession: S15132  
A;Molecule type: protein  
A;Residues: 29-38;76-80 <BIO>  
A;Note: forms RSC-1, RSC-2 and RSC-3 with differing amino-terminals were found; a form be  
C;Genetics.  
A;Gene: Cys8  
A;Introns: 76/3; 114/3  
C;Superfamily: cystatin; cystatin homology  
C;Keywords: cysteine proteinase inhibitor  
F;1-28/Domain: signal sequence #status experimental <SIG>  
F;28-141/Product: cystatin S #status experimental <MAT>  
F;29-141/Product: cystatin S, form RSC-3 #status predicted <MAT3>  
F;30-141/Domain: cystatin homology <CYS>  
F;31-141/Product: cystatin S, form RSC-2 #status predicted <MAT2>  
F;32-141/Product: cystatin S, form RSC-1 #status predicted <MAT1>  
F;94-104,118-138/Disulfide bonds: #status experimental

Query Match 20.6%; Score 129.5; DB 2; Length 141;  
Best Local Similarity 31.7%; Pred. NO. 7.8e-06;  
Matches 32; Conservative 20; Mismatches 44; Indels 5; Gaps 3;  
Qy 18 ENYAKDSLQMTTDOYNKESDDKHFRIPRLVKVQSQVTDHLBYHLNVEMQWTTCCQKP--E 75  
Db 41 EEGASEALNVAVNEYKNKSDLYLSRVVEVDKQVAGTKFFFDVILGKTLCTQGD 100  
Qy 76 TTNCVQPERLHKQVN--CPFSFVAPMPFEQYKILNKSCSS 114  
Db 101 LTNC-PLNEEADQEHCFEBCFSFVHDIPWENYVILLSSSCHS 140

RESULT 14  
A28793  
Cystatin - puff adder  
C;Species: Bitis arietans (puff adder)  
C;Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 30-Sep-1993  
C;Accession: A28793  
R;Ritonja, J.; Evans, H.J.; Machleidt, W.; Barrett, A.J.  
Biochem. J. 246, 799-802, 1987  
A;Title: Amino acid sequence of a cystatin from venom of the African puff adder (Bitis a  
A;Reference number: A28793; MUID:88076861; PMID:3500714  
A;Accession: A28793  
A;Molecule type: protein  
A;Residues: 1-111 <RIT>  
C;Superfamily: cystatin; cystatin homology

Cysteine 3 precursor rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text change 16-Jul-1999

F;Ritonja, A.; Evans, H.J.; Machleidt, W.; Barrett, A.J.  
Biochem. J. 246, 799-802, 1987  
A;Title: Amino acid sequence of a cystatin from venom of  
A;Reference number: A28793; MUID:88076861; PMID:3500714

A:Cross-references: GB:J03870; NID:g337751; PIDN:AAA60299.1; PID:g337752  
R:Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.  
Biol. Chem. Hoppe-Seyler 369, 191-197, 1988  
A:Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily  
A:Reference number: S02489; MUID:89076505; PMID:3202964  
A:Accession: S02489  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 21-141 <SA2>  
R:Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.  
Gene 61, 329-338, 1987  
A:Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three mem  
A:Reference number: A91589; MUID:88185836; PMID:3446578  
A:Accession: A29632  
A:Molecule type: DNA  
A:Residues: 1-86, 'I', 88-141 <SAI>  
R:Isemura, S.; Saitoh, E.; Sanada, K.  
FEBS Lett. 198, 145-149, 1986  
A:Title: Characterization of a new cysteine proteinase inhibitor of human saliva, cystat  
A:Reference number: A01273; MUID:86164938; PMID:3514272  
A:Accession: A01273  
A:Molecule type: protein  
A:Residues: 29-141 <ISE>  
R:Ramakrishnan, N.; Reddy, M.S.; Bergey, E.J.; Haraszthy, G.G.; Soni, S.D.; Levine, M.J.  
Biochem. J. 280, 341-352, 1991  
A:Title: Large-scale purification and characterization of the major phosphoproteins and  
A:Reference number: S19279; MUID:92082469; PMID:1747107  
A:Accession: S19279  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 21-55 <RAM>  
C:Comment: Human saliva appears to contain several cysteine proteinase inhibitors that an  
ences. Cystatin SN, with a pi of 7.5, is a much better inhibitor of papain and dipeptidyl  
C:Genetics:  
A:Gene: GDB:CST1  
A:Cross-references: GDB:119815; OMIM:123855  
A:Map position: 20p11.2-20p11.2  
C:Superfamily: cystatin; cystatin homology  
C:Keywords: cysteine proteinase inhibitor; extracellular protein; saliva  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-141/Product: cystatin SA-I #status experimental <MAT1>  
F:29-141/Product: cystatin SN #status experimental <MAT2>  
F:30-141/Domain: cystatin homology <CVS>  
F:76-80/Region: inhibitory #status predicted  
F:94-104,118-138/Disulfide bonds: #status predicted

Query Match 22.2%; Score 139.5; DB 1; Length 141;  
Best Local Similarity 29.6%; Pred. No. 7.9e-07;  
Matches 29; Conservative 21; Mismatches 45; Indels 3; Gaps 2;

Qy 18 ENYAKDSLQWITDQVKNESDKYFIRFLVKVQRQVTDHLEYHLNVEMQWTTCKQ--PE 75  
Db 41 DEWQRALHFAISEYNKATKODYRNPRLVRLARQQTGGVNYFFDEVGRICTKSQPN 100  
Qy 76 TTNCVQPER-ELHKQVNCFFSVFVAPWPEQYKILNKSC 112  
Db 101 LDTCAFHEQPELQKQLCSFEIYEVPWENRRSLVKSRC 138

RESULT 11  
JC2040  
cystatin - chum salmon  
N:Alternate names: cysteine proteinase inhibitor  
C:Species: Oncorhynchus keta (chum salmon)  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 16-Jul-1999  
C:Accession: JC2040  
R:Koide, Y.; Noso, T.  
BioSci. Biotechnol. Biochem. 58, 164-169, 1994  
A:Title: The complete amino acid sequence of pituitary cystatin from chum salmon.  
A:Reference number: JC2040; MUID:94162738; PMID:7764512  
A:Accession: JC2040  
A:Molecule type: protein  
A:Residues: 1-111 <KOI>

A;Accession: A01270  
A;Molecule type: protein  
A;Residues: 27-131,'S',133-146 <GRU>  
R;Chisio, J.; Jenson, O.; Frangione, B.  
Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986  
A;Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland  
A;Reference number: A25434; MUID:86206076; PMID:3517880  
A;Accession: A25434  
A;Molecule type: protein  
A;Residues: 37-93,'O',95-146 <GHI>  
R;Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983  
A;Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin  
A;Reference number: S01461; MUID:84110059; PMID:6662498  
A;Accession: S12288  
A;Molecule type: protein  
A;Residues: 27-73 <TUR>  
R;Brzin, J.; Popovic, T.; Turk, V.  
Biochem. Biophys. Res. Commun. 118, 103-109, 1984  
A;Title: Human cystatin, a new protein inhibitor of cysteine proteinases.  
A;Reference number: A32732; MUID:84128015; PMID:6365094  
A;Accession: A32732  
A;Molecule type: protein  
A;Residues: 27-76 <BRZ>  
R;Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jenson, O.; Grubb, A.  
Scand. J. Clin. Lab. Invest. 50, 85-93, 1990  
A;Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst  
A;Reference number: A60552; MUID:90193615; PMID:2315647  
A;Accession: A60552  
A;Molecule type: protein  
A;Residues: 27-49,'XX',52-64 <OLA>  
A;Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d  
e defective gene is not present in CSF but is found instead in amyloid deposits  
R;Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.  
Biol. Chem. Hoppe-Seyler 371, 575-580, 1990  
A;Title: Different forms of human cystatin C.  
A;Reference number: S10607; MUID:91025625; PMID:2222856  
A;Accession: S10607  
A;Molecule type: protein  
A;Residues: 27-53 <POP>  
A;Experimental source: urine, kidney disease  
A;Note: truncated forms with amino ends at positions 35 and 36 of the precursor were als  
R;Grubb, A.; Lofberg, H.; Barrett, A.J.  
FEBS Lett. 170, 370-374, 1984  
A;Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
A;Reference number: S01462  
A;Contents: annotation; disulfide bonds  
R;Berti, P.J.; Storer, A.C.  
Biochem. J. 302, 411-416, 1994  
A;Title: Local pH-dependent conformational changes leading to proteolytic susceptibility  
A;Reference number: S55305; MUID:94379969; PMID:8092991  
A;Accession: S55305  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 27-49;106-146 <BER>  
C;Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl  
F patients with certain autoimmune diseases.  
C;Comment: This protein is an inhibitor of cysteine proteinases and may serve an importa  
C;Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorr  
C;Genetics:  
A;Gene: GDB:CST3  
A;Cross-references: GDB:119817; OMIM:105150  
A;Map position: 20p11.2-20p11.2  
A;Introns: 81/3; 119/3  
C;Superfamily: cystatin; cystatin homology  
C;Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyprolin  
F;1-26/Domain: signal sequence #status predicted <SIG>  
F;27-146/Product: cystatin C #status experimental <MAT>  
F;35-146/Domain: cystatin homology <CYS>  
F;81-85/Region: inhibitory #status predicted  
F;29/Modified site: hydroxyproline (Pro) (partial) #status experimental  
F;99-109,123-143/Disulfide bonds: #status experimental

Query Match 24.6%; Score 155; DB 1; Length 146;  
Best Local Similarity 33.7%; Pred. No. 2.3e-08;  
Matches 34; Conservative 21; Mismatches 42; Indels 4; Gaps 3;  
QY 18 ENYAKDSLOWITDQYNKESDDKYHFRIVLKVQVQVTDHLEYHLNVEMQWTTTCQK--PE 75  
DB 46 EGVRRALDPANVEYNKASNDMYHSALQVVRARKQIVAGVNYFLDVLGRITTCCTQPN 105  
QY 76 TTNC-VPOREHLHKQVNCFFSVFAVWFQYKILNKSCSSD 115  
DB 106 LDNCPPHDQPHLKRKAFCSQIYAVDW-QGTWTLSTKSTQD 145  
RESULT 8  
UDROP  
cystatin - bovine  
N;Alternate names: thiol proteinase inhibitor  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 28-Feb-1986 #sequence\_revision 28-Feb-1986 #text\_change 06-Dec-1996  
C;Accession: A01271  
R;Hirado, M.; Tsunagawa, S.; Sakiyama, F.; Niinobe, M.; Fujii, S.  
FEBS Lett. 186, 41-45, 1985  
A;Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase in  
A;Reference number: A01271; MUID:85231205; PMID:3891407  
A;Accession: A01271  
A;Molecule type: protein  
A;Residues: 1-112 <HIR>  
C;Superfamily: cystatin; cystatin homology  
C;Keywords: colostrum; cysteine proteinase inhibitor  
F;2-112/Domain: cystatin homology <CYS>  
F;48-52/Region: inhibitory #status predicted  
F;66-76,90-110/Disulfide bonds: #status predicted  
Query Match 24.4%; Score 153.5; DB 1; Length 112;  
Best Local Similarity 30.6%; Pred. No. 2.4e-08;  
Matches 30; Conservative 23; Mismatches 42; Indels 3; Gaps 2;  
QY 18 ENYAKDSLOWITDQYNKESDDKYHFRIVLKVQVQVTDHLEYHLNVEMQWTTTCQKPE 77  
DB 13 EGVQEQALSFVAVSEFNKRNDAVQSRVVRVVRKQVSGMNYFLDVLGRITTCCTQSAN 72  
QY 78 --NC-VPOREHLHKQVNCFFSVFAVWFQYKILNKSC 112  
DB 73 LDSCPPHNPFLKREKLCFQVYVVPVMTINLVKESC 110  
RESULT 9  
UDROP1  
cystatin S precursor - human  
N;Alternate names: cystatin SA-III; salivary acidic protein-1  
C;Species: Homo sapiens (man)  
C;Date: 25-Feb-1985 #sequence\_revision 08-Feb-1996 #text\_change 16-Jul-1999  
C;Accession: S17667; S16500; A01272; A29603; S19280; A56508  
R;Bobek, L.A.; Aguirre, A.; Levine, M.J.  
Biochem. J. 278, 629-635, 1991  
A;Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ an  
A;Reference number: S17667; MUID:91378918; PMID:1898352  
A;Accession: S17667  
A;Molecule type: mRNA  
A;Residues: 1-141 <SOB>  
A;Cross-references: EMBL:X54667; NID:G30365; PIDN:CAA38478.1; PID:G30366  
R;Lamkin, M.S.; Jensen, J.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.  
Arch. Biochem. Biophys. 288, 664-670, 1991  
A;Title: Salivary cystatin SA-III, a potential precursor of the acquired enamel pellicle,  
A;Reference number: S16500; MUID:91378515; PMID:1898055  
A;Accession: S16500  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 21-134,'D',136-141 <IHU>  
R;Isemura, S.; Saitoh, E.; Sanada, K.  
J. Biochem. 96, 489-498, 1984  
A;Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole sal  
A;Reference number: A91985; MUID:85054716; PMID:6501254





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 22.1329 Seconds  
(without alignments)  
499.799 Million cell updates/sec

Title: US-09-941-314-3  
Perfect score: 629  
Sequence: 1 YQARKKTFLSVHEWAVENY.....VFAVPWFQYKILNKSCSD 115

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues 283366  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: P1r1:\*  
2: P1r2:\*  
3: P1r3:\*  
4: P1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214	34.0	139	2 A45361	cystatin-related e
2	171.5	27.3	127	2 S07085	cystatin C precurs
3	170.5	27.1	139	1 UDCH	cystatin precursor
4	167.5	26.6	120	2 S10587	cystatin C - rat
5	160.5	25.5	141	2 B29632	cystatin SA precur
6	156.5	24.9	140	2 A36163	cystatin C precurs
7	155	24.6	146	1 UDHU	cystatin C precurs
8	153.5	24.4	112	1 UDHO	cystatin - bovine
9	146.5	23.3	141	1 UDHUP1	cystatin S precurs
10	139.5	22.2	141	1 UDHUP2	cystatin SN precur
11	134	21.3	111	1 JC2040	cystatin - chum sa
12	131.5	20.9	142	2 A47442	cystatin D precurs
13	129.5	20.6	141	2 JQ1470	cystatin S precurs
14	127	20.2	111	2 A28793	cystatin - puff ad
15	126	20.0	132	2 JC4918	cystatin precursor
16	102.5	16.3	133	2 JC4536	cystatin precursor
17	98.5	15.7	434	1 KGBOL2	kininogen, LMW II
18	98.5	15.7	639	1 KGBOH2	kininogen, HMW II
19	96.5	15.3	436	1 KGBOL1	kininogen, LMW I p
20	96.5	15.3	621	1 KGBOH1	kininogen, HMW I p
21	95.5	15.2	162	2 A43428	onchocystatin - ne
22	93.5	14.9	427	1 KGHUL1	kininogen, LMW pre
23	93.5	14.9	644	1 KGHUL1	kininogen, HMW pre
24	87	13.8	433	2 A28055	X-kininogen, LMW I
25	87	13.8	639	2 A25486	kininogen, HMW I p
26	81	12.9	430	2 A23897	major acute phase
27	81	12.9	430	2 B28055	T-kininogen, LMW I
28	80	12.7	498	2 T31871	hypothetical prote
29	79	12.6	430	1 KGRITI	T-kininogen I prec

30	77	12.2	430	2 D64151	hypothetical prote
31	75	11.9	400	2 A46297	beta-1,6-N-acetylgl
32	75	11.9	423	1 KGRIM	major acute phase
33	74	11.8	582	2 S42613	membrane protein p
34	73.5	11.7	438	2 A47702	glucan 1,3-beta-gl
35	73.5	11.7	438	2 T52149	beta-glucanase lim
36	73	11.6	861	2 S12499	CHLI protein - yea
37	71.5	11.4	132	2 D88508	protein H14A12.5 (
38	71.5	11.4	1227	2 T23004	hypothetical prote
39	71.5	11.4	1779	2 T23130	hypothetical prote
40	71	11.3	677	1 SYECMT	methionine-tRNA 11
41	71	11.3	677	2 C85839	methionine tRNA sy
42	71	11.3	677	2 H90993	methionine tRNA sy
43	70.5	11.2	448	2 JN0118	glucan 1,3-beta-gl
44	70	11.1	184	2 D72061	hypothetical prote
45	70	11.1	184	2 B86562	hypothetical prote

ALIGNMENTS

RESULT 1

A45361  
cystatin-related epididymal specific protein - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C;Accession: A45361  
R;Corrwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.  
Mol. Endocrinol. 6, 1653-1664, 1992  
A;Title: The CRIS gene: a unique testis-regulated gene related to the cystatin family is  
A;Reference number: A45361, MUID:93078799; PMID:1280328  
A;Accession: A45361  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: nucleic acid  
A;Residues: 1-139 <OR>  
A;Cross-references: GB:S49926; NID:G260492; PIDN:AAC35390.1; PID:G260493  
A;Note: sequence extracted from NCBI backbone (NCBIP:118813)  
C;Superfamily: cystatin; cystatin homology  
F;28-139/Domain: cystatin homology <CYS>

Query Match	34.0%;	Score 214;	DB 2;	Length 139;
Best Local Similarity	37.1%;	Pred. NO. 2.9e-14;		
Matches	43;	Conservative 29;	Mismatches 30;	Indels 14; Gaps 4;
QY	10	SVHEWAVENY-----AKDSLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLE	59	
DB	22	SKNEVKA-QNYFGSINISNANVKQCVWFAMKYNKESDKYFLVDKILHAKLQITDRME	80	
QY	60	YHLNVEMQWTTCKP--ETTNCVPQER-ELHKOVNCFVSFVAFVPWFQYKILNKSC	112	
DB	81	YQIDVQISRSNCKPLNNTENCIPQKKPELEKMKSCFLVGLPWFNGEFNLLSKEC	136	

RESULT 2

S07085  
cystatin C precursor - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 01-Dec-1993 #sequence\_revision 03-Aug-1995 #text\_change 16-Jul-1999  
C;Accession: S07085; S01337; S21109  
R;Cole, T.; Dickson, P.W.; Esmard, F.; Averill, S.; Risbridger, G.P.; Gauthier, F.; Schre  
Eur. J. Biochem. 186, 35-42, 1989  
A;Title: The cDNA structure and expression analysis of the genes for the cysteine protei  
A;Reference number: S07085; MUID:90092122; PMID:2689174  
A;Accession: S07085  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-127 <COL>  
A;Cross-references: EMBL:X16957; NID:G56041; PIDN:CAA34831.1; PID:G736290  
R;Esmard, A.; Esmard, F.; Faucher, D.; Gauthier, F.  
FEBS Lett. 236, 475-478, 1988  
A;Title: Two rat homologues of human cystatin C.  
A;Reference number: S01337; MUID:88313020; PMID:3044831  
A;Accession: S01337



; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-6

Query Match 30.0%; Score 189; DB 9; Length 36;  
Best Local Similarity 100.0%; Pred. NO. 3.6e-14;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 RKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 39  
Db 1 RKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36

RESULT 14  
US-09-941-314-8  
; Sequence 8, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-8

Query Match 29.7%; Score 187; DB 9; Length 35;  
Best Local Similarity 100.0%; Pred. NO. 5.9e-14;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 22 KDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTD 56  
Db 1 KDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTD 35

RESULT 15  
US-09-740-638-2  
; Sequence 2, Application US/09740638  
; Patent No. US20020006856A1  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; TITLE OF INVENTION: 2cys5: A Member of the Cystatin  
; TITLE OF INVENTION: Superfamily  
; FILE REFERENCE: 99-104  
; CURRENT APPLICATION NUMBER: US/09/740,638  
; CURRENT FILING DATE: 2000-12-18  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 145  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-740-638-2

Query Match 28.2%; Score 177.5; DB 9; Length 145;

Best Local Similarity 31.4%; Pred. No. 4e-12;  
Matches 32; Conservative 30; Mismatches 37; Indels 3; Gaps 2;  
QY 13 EYMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEHQWTTCC 72  
Db 34 KLMSKKK-MNSTLNFFIQSYNNASNDTYLYRVQRLIRSQRLTGVGEYIVTVKIGWTKCK 92  
QY 73 KPETT--CVQPERLHKOVNCFPSVFAVPWFQYKILNKSC 112  
Db 93 RNDTSNSSCPLOSKKURKSLICESLIYTWEPWYNYFQLWNNSC 134

Search completed: March 18, 2004, 14:30:52  
Job time : 70.5933 secs

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 53  
 Db 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 10  
 US-09-864-761-48936  
 ; Sequence 48936, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 ; FILE REFERENCE: Aeomica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864, 761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 48936  
 ; LENGTH: 50  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AL096677.18  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96  
 ; OTHER INFORMATION: EST HUMAN HIT: AI200857.1, EVALUATE 5.00e-23  
 ; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01

US-09-864-761-48936  
 ; Sequence 48936, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 ; FILE REFERENCE: Aeomica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864, 761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 48936  
 ; LENGTH: 50  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AL096677.18  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96  
 ; OTHER INFORMATION: EST HUMAN HIT: AI200857.1, EVALUATE 5.00e-23  
 ; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01

Query Match 41.7%; Score 262; DB 9; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 53

Db 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 11  
 US-09-941-314-10  
 ; Sequence 10, Application US/09941314  
 ; Patent No. US20020142396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZymoGenetics, Inc.  
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
 ; FILE REFERENCE: 00-81PC  
 ; CURRENT APPLICATION NUMBER: US/09/941,314  
 ; CURRENT FILING DATE: 2001-08-29  
 ; PRIOR APPLICATION NUMBER: 60/230,230  
 ; PRIOR FILING DATE: 2001-09-01  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 10  
 ; LENGTH: 46  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-941-314-10

Query Match 40.4%; Score 254; DB 9; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-21;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCKPPTTN 78  
 Db 1 NKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCKPPTTN 46

RESULT 12  
 US-09-941-314-12  
 ; Sequence 12, Application US/09941314  
 ; Patent No. US20020142396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZymoGenetics, Inc.  
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
 ; FILE REFERENCE: 00-81PC  
 ; CURRENT APPLICATION NUMBER: US/09/941,314  
 ; CURRENT FILING DATE: 2001-08-29  
 ; PRIOR APPLICATION NUMBER: 60/230,230  
 ; PRIOR FILING DATE: 2001-09-01  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 12  
 ; LENGTH: 33  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-941-314-12

Query Match 30.0%; Score 189; DB 9; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-14;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 QVTDHLEYHLNVEMQWTTCKPPTTNQVQERE 85  
 Db 1 QVTDHLEYHLNVEMQWTTCKPPTTNQVQERE 33

RESULT 13  
 US-09-941-314-6  
 ; Sequence 6, Application US/09941314  
 ; Patent No. US20020142396A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZymoGenetics, Inc.  
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to

```

; ORGANISM: Homo sapiens
US-09-941-314-14
Query Match      45.8%; Score 288; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 3.4e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 KESDDKYHFRFVLKQVQVTDHLEHLNVNQWTTCKPPTNCVPOERE 85
      |||
      1 KESDDKYHFRFVLKQVQVTDHLEHLNVNQWTTCKPPTNCVPOERE 52

RESULT 7
US-09-941-314-17
; Sequence 17, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-17

Query Match      43.4%; Score 273; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 NVEMQWTTCKPPTNCVPOERHLKQVNCFFSVFVPEQYKILNK 110
      |||
      1 NVEMQWTTCKPPTNCVPOERHLKQVNCFFSVFVPEQYKILNK 48

RESULT 8
US-09-941-314-13
; Sequence 13, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-13

Query Match      43.2%; Score 272; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.1e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 DQYNKESDDKYHFRFVLKQVQVTDHLEHLNVNQWTTCKPPTTN 78
      |||
      1 DQYNKESDDKYHFRFVLKQVQVTDHLEHLNVNQWTTCKPPTTN 49

RESULT 9
US-09-864-761-34822
; Sequence 34822, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34822
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109954.10
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
US-09-864-761-34822

Query Match      41.7%; Score 262; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;

```

```

; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-4

Query Match      100.0%; Score 629; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.1e-63;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQVQVTDHLEY 60
DB 3 YQARKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQVQVTDHLEY 62

QY 61 HLNVEQWTTTCQKPEPTTNCVQRELHKQVNCFFSVFVFPWFQYKILNKSCSSD 115
DB 63 HLNVEQWTTTCQKPEPTTNCVQRELHKQVNCFFSVFVFPWFQYKILNKSCSSD 117

RESULT 3
US-09-941-314-2
; Sequence 2, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-2

Query Match      100.0%; Score 629; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.6e-63;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQVQVTDHLEY 60
DB 23 YQARKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQVQVTDHLEY 82

QY 61 HLNVEQWTTTCQKPEPTTNCVQRELHKQVNCFFSVFVFPWFQYKILNKSCSSD 115
DB 83 HLNVEQWTTTCQKPEPTTNCVQRELHKQVNCFFSVFVFPWFQYKILNKSCSSD 137

RESULT 4
US-09-941-314-15
; Sequence 15, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC

```

```

; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-15

Query Match      70.9%; Score 446; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 7.3e-43;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 QYNKESDDKYHFRIFRLVKVQVQVTDHLEYHLNVEMQWTTTCQKPEPTTNCVQRELHKQV 90
DB 1 QYNKESDDKYHFRIFRLVKVQVQVTDHLEYHLNVEMQWTTTCQKPEPTTNCVQRELHKQV 60

QY 91 NCFFSVFVFPWFQYKILNK 110
DB 61 NCFFSVFVFPWFQYKILNK 80

RESULT 5
US-09-941-314-16
; Sequence 16, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-16

Query Match      53.1%; Score 334; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.5e-30;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 RQVTDHLEYHLNVEMQWTTTCQKPEPTTNCVQRELHKQVNCFFSVFVFPWFQYKILNK 110
DB 1 RQVTDHLEYHLNVEMQWTTTCQKPEPTTNCVQRELHKQVNCFFSVFVFPWFQYKILNK 59

RESULT 6
US-09-941-314-14
; Sequence 14, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 52
; TYPE: PRT

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 69.5913 Seconds  
(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-3  
Perfect score: 629  
Sequence: 1 YQARKKTFSLVHEVMAVENY.....VFAPWFQYKILNKSCSSD 115

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pap.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pap.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pap.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pap.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pap.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pap.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pap.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pap.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pap.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pap.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pap.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pap.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	629	100.0	115	9 US-09-941-314-3	Sequence 3, Appli
2	629	100.0	117	9 US-09-941-314-4	Sequence 4, Appli
3	629	100.0	117	9 US-09-941-314-2	Sequence 2, Appli
4	446	70.9	80	9 US-09-941-314-15	Sequence 15, Appl
5	334	53.1	59	9 US-09-941-314-16	Sequence 16, Appl
6	288	45.8	52	9 US-09-941-314-14	Sequence 14, Appl
7	273	43.2	48	9 US-09-941-314-17	Sequence 17, Appl
8	272	43.2	49	9 US-09-941-314-13	Sequence 13, Appl
9	262	41.7	50	9 US-09-864-761-14822	Sequence 34822, A
10	262	41.7	50	9 US-09-864-761-48936	Sequence 48936, A
11	254	40.4	46	9 US-09-941-314-10	Sequence 10, Appl
12	189	30.0	33	9 US-09-941-314-12	Sequence 12, Appl
13	189	30.0	36	9 US-09-941-314-6	Sequence 6, Appli
14	187	29.7	35	9 US-09-941-314-8	Sequence 8, Appli
15	177.5	28.2	145	9 US-09-740-638-2	Sequence 2, Appli

16	177.5	28.2	145	13	US-10-006-467-2	Sequence 2, Appli
17	177.5	28.2	145	14	US-10-235-148-2	Sequence 2, Appli
18	177.5	27.3	145	8	US-08-849-303-19	Sequence 19, Appl
19	171.5	27.3	145	14	US-10-168-425-14	Sequence 14, Appl
20	170.5	27.1	116	9	US-09-775-932-16	Sequence 16, Appl
21	170.5	27.1	116	9	US-08-849-303-15	Sequence 15, Appl
22	170.5	27.1	139	9	US-09-969-834-4	Sequence 4, Appli
23	160.5	25.5	121	9	US-09-775-932-8	Sequence 8, Appli
24	160.5	25.5	140	14	US-10-376-564-46	Sequence 46, Appl
25	160.5	25.5	141	8	US-08-849-303-24	Sequence 24, Appl
26	160.5	25.5	141	9	US-09-940-497-6	Sequence 6, Appli
27	159.5	25.4	181	15	US-10-264-049-2608	Sequence 2608, Ap
28	157.5	25.0	165	13	US-09-740-638-5	Sequence 5, Appli
29	157.5	25.0	165	13	US-10-006-467-5	Sequence 5, Appli
30	157.5	25.0	165	14	US-10-235-148-5	Sequence 5, Appli
31	156.5	24.9	140	8	US-08-849-303-18	Sequence 18, Appl
32	156.5	24.9	140	14	US-10-376-564-48	Sequence 48, Appl
33	155	24.6	120	9	US-09-775-932-2	Sequence 2, Appli
34	155	24.6	146	8	US-08-849-303-17	Sequence 17, Appl
35	155	24.6	146	9	US-09-940-497-3	Sequence 3, Appli
36	155	24.6	146	9	US-09-969-834-3	Sequence 3, Appli
37	155	24.6	146	14	US-10-329-428-3	Sequence 3, Appli
38	155	24.6	146	14	US-10-376-564-47	Sequence 47, Appl
39	153.5	24.4	112	8	US-08-849-303-16	Sequence 16, Appl
40	153.5	24.4	118	9	US-09-775-932-24	Sequence 24, Appl
41	146.5	23.3	121	9	US-09-775-932-4	Sequence 4, Appli
42	146.5	23.3	141	8	US-08-849-303-22	Sequence 22, Appl
43	146.5	23.3	141	9	US-09-940-497-5	Sequence 5, Appli
44	146.5	23.3	141	9	US-09-974-298-141	Sequence 141, App
45	146.5	23.3	141	14	US-10-241-220-77	Sequence 77, Appl

ALIGNMENTS

RESULT 1  
US-09-941-314-3  
; Sequence 3, Application US/09941314  
; Patent No. US20020142396A1  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
; FILE REFERENCE: 00-81PC  
; CURRENT APPLICATION NUMBER: US/09/941,314  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/230,230  
; PRIOR FILING DATE: 2001-09-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-314-3

Query Match 100.0%; Score 629; DB 9; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.1e-63;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	YQARKKTFSLVHEVMAVENYAKDSLOWITDQYNKSDDKYHFRIFRVLKQVQRTDILEY	60
Db	1	YQARKKTFSLVHEVMAVENYAKDSLOWITDQYNKSDDKYHFRIFRVLKQVQRTDILEY	60
Qy	61	HLNVEMQWTTCCQPETTNCVQPERLHKQVNCFFSFAVFWFQYKILNKSCSSD	115
Db	61	HLNVEMQWTTCCQPETTNCVQPERLHKQVNCFFSFAVFWFQYKILNKSCSSD	115

RESULT 2  
US-09-941-314-4  
; Sequence 4, Application US/09941314  
; Patent No. US20020142396A1

```

; TELEFAX: 301 309 8512
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Cystatin SN
; US-08-744-138-6

Query Match      25.5%; Score 160.5; DB 3; Length 141;
Best Local Similarity 31.5%; Pred. No. 6.2e-12;
Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;

QY      24 SLOWITDQYNKESDDKYHFRIFPVLVKQVQVTDHLEYHLNVEMQWTTCK--PETTNCVP 81
Db      47 ALHFVISEYNKATEDEYRRLLEVLRAREQIVGGVNYFFDIEVGRITCTKSPNLDTCAP 106

QY      82 QER-ELHKQVNCFFSVFAVPWFQYKILNKSC 112
Db      107 HEQPELQKKQLCSFQIYEVPEWEDMSLVNSRC 138

```

Search completed: March 18, 2004, 14:25:34  
Job time : 30.4345 secs

RESULT 14

```
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0193 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 139 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 118195
/ US-08-791-522-4

Query Match 27.1%; Score 170.5; DB 2; Length 139;
Best Local Similarity 37.0%; Pred. No. 3.6e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 24 SLOWITDQYNKESDDKYHFRIFRVLKVRQVTDHLYHLNVEMQWTTCKP--ETNC-V 80
Db 47 ALQFAMAEYNRASNDKYSRRVVRVISAKRLVSGIKYILQVEIGRTTCPKSSGDLOSCEF 106

QY 81 PQERELHKQVNCFFSVFVFWPEQYKILNKC 112
Db 107 HDEPEMAKYTTCTFVVYSIPWLNQIKLESKC 138

RESULT 10
US-09-314-777-4
/ Sequence 4, Application US/09314777
/ Patent No. 6110686
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Goli, Surya K.
/ TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
/ TITLE OF INVENTION: PROTEIN
/ NUMBER OF SEQUENCES: 4
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/314,777
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/791,522
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0193 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 139 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
```

```
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 118195
/ US-09-314-777-4

Query Match 27.1%; Score 170.5; DB 3; Length 139;
Best Local Similarity 37.0%; Pred. No. 3.6e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 24 SLOWITDQYNKESDDKYHFRIFRVLKVRQVTDHLYHLNVEMQWTTCKP--ETNC-V 80
Db 47 ALQFAMAEYNRASNDKYSRRVVRVISAKRLVSGIKYILQVEIGRTTCPKSSGDLOSCEF 106

QY 81 PQERELHKQVNCFFSVFVFWPEQYKILNKC 112
Db 107 HDEPEMAKYTTCTFVVYSIPWLNQIKLESKC 138

RESULT 11
US-08-849-303-15
/ Sequence 15, Application US/08849303
/ Patent No. 6680424
/ GENERAL INFORMATION:
/ APPLICANT: Atkinson, Howard J.
/ APPLICANT: McPherson, Michael J.
/ APPLICANT: Urwin, Peter E.
/ TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
/ NUMBER OF SEQUENCES: 79
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Klauber & Jackson
/ STREET: 411 Hackensack Avenue, 4th Floor
/ CITY: Hackensack
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07601
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/849,303
/ FILING DATE: 21-MAY-1997
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Jackson Esq., David A.
/ REGISTRATION NUMBER: 26,742
/ REFERENCE/DOCKET NUMBER: 1321-1-003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 201-487-5800
/ TELEFAX: 201-343-1684
/ TELEX: 133521
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 139 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ US-08-849-303-15

Query Match 27.1%; Score 170.5; DB 4; Length 139;
Best Local Similarity 37.0%; Pred. No. 3.6e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 24 SLOWITDQYNKESDDKYHFRIFRVLKVRQVTDHLYHLNVEMQWTTCKP--ETNC-V 80
Db 47 ALQFAMAEYNRASNDKYSRRVVRVISAKRLVSGIKYILQVEIGRTTCPKSSGDLOSCEF 106

QY 81 PQERELHKQVNCFFSVFVFWPEQYKILNKC 112
Db 107 HDEPEMAKYTTCTFVVYSIPWLNQIKLESKC 138
```

```

; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-2

Query Match      29.8%; Score 187.5; DB 3; Length 141;
Best Local Similarity 39.6%; Pred. No. 3.1e-15;
Matches 38; Conservative 20; Mismatches 35; Indels 3; Gaps 2;

QY 20 YAKDSLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTQK--PETT 77
Db 43 YVQHALWYAMKEYNKASNDLYNFRVVDILKSQEQITDSLEYLYEVNIARTMCKKIAGDNE 102

QY 78 NCV-PQERELHKQVNCFFSVFVFWPEQYKILNKC 112
Db 103 NCLFQDDPKMKVMFCIFIVSSKPKWKFELKMLKKQC 138

RESULT 7
US-08-849-303-19
; Sequence 19, Application US/08849303
; Patent No. 6680424
; GENERAL INFORMATION:
; APPLICANT: Atkinson, Howard J.
; APPLICANT: McPherson, Michael J.
; APPLICANT: Urwin, Peter E.
; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,303
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-849-303-19

Query Match      27.3%; Score 171.5; DB 4; Length 127;
Best Local Similarity 33.3%; Pred. No. 2.4e-13;
Matches 35; Conservative 24; Mismatches 43; Indels 3; Gaps 2;

```

```

QY 13 EYMAVENAYKDSLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTQK 72
Db 22 EADASEGVQALDPAVSEYNGSDAYHSRAIQVVRARQLVAGINYLDVEMGRITCT 81

QY 73 KPET--TNC-VQERELHKQVNCFFSVFVFWPEQYKILNKCSCS 114
Db 82 KSQTNLTNCPFDQPHLMRKALCSFOIYSVPWKGTHLTLSCKN 126

RESULT 8
US-09-775-932-16
; Sequence 16, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cystatins
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-775-932-16

Query Match      27.1%; Score 170.5; DB 4; Length 116;
Best Local Similarity 37.0%; Pred. No. 2.9e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 24 SLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTQK--ETNC-V 80
Db 24 ALQFAMAENYRASNDKYSSRVVVISAKQLVSGIKYILQVEIGRTTCPSKSDQLQSCF 83

QY 81 PQERELHKQVNCFFSVFVFWPEQYKILNKC 112
Db 84 HDEPEMAKYTCTCTFVVSIPMLNQLKLSKC 115

RESULT 9
US-08-791-522-4
; Sequence 4, Application US/08791522
; Patent No. 5935817
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,522
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

; PRIOR FILING DATE: 1998-11-20  
; PRIOR APPLICATION NUMBER: 60/156,382  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-617-302-3

Query Match 34.0%; Score 214; DB 3; Length 142;  
Best Local Similarity 37.1%; Pred. No. 1.8e-18;  
Matches 43; Conservative 29; Mismatches 30; Indels 14; Gaps 4;  
QY 10 SVHEVMAVENY-----AKDSLQWITDQYNKESDDKYHFRIPRLVKVQVQVTDHLE 59  
DB 25 SKNEVKA-QNYFGSINISNANVCQVFWFAMKEYNKESDKYFLVDKILHAKLQITDRME 83  
QY 60 YHLNVEMQWTTCKP--ETTNCVPQER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 112  
DB 84 YQIDVQISRSNCKKPLNNTENCIPQKPELEKMKSCSFLVGALPWNGEFNLKSEC 139

RESULT 3  
US-09-431-480-4  
; Sequence 4, Application US/09431480  
; Patent No. 6235708  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72  
; CURRENT APPLICATION NUMBER: US/09/431,480  
; CURRENT FILING DATE: 1999-11-01  
; EARLIER APPLICATION NUMBER: 60/109,217  
; EARLIER FILING DATE: 1998-11-20  
; EARLIER APPLICATION NUMBER: 60/156,382  
; EARLIER FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-431-480-4

Query Match 30.0%; Score 188.5; DB 3; Length 142;  
Best Local Similarity 40.4%; Pred. No. 2.4e-15;  
Matches 38; Conservative 23; Mismatches 30; Indels 3; Gaps 2;  
QY 22 KDSLQWITDQYNKESDDKYHFRIPRLVKVQVQVTDHLEHLNVEMQWTTCKPPTTN--C 79  
DB 46 KQCLWFAMQYNEKSEDKYFVLVVKTLQALQVNTNLEVLIDVEIARSDCRKPLSTNEIC 105  
QY 80 VPQER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 112  
DB 106 AIQENSKLKRKLSCLSGFLVGALPWNGEFTWMEKCC 139

RESULT 4  
US-09-617-302-4  
; Sequence 4, Application US/09617302  
; Patent No. 6245529  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72 C1  
; CURRENT APPLICATION NUMBER: US/09/617,302  
; CURRENT FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/431,480  
; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: 60/109,217  
; PRIOR FILING DATE: 1998-11-20  
; PRIOR APPLICATION NUMBER: 60/156,382  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-617-302-4

Query Match 30.0%; Score 188.5; DB 3; Length 142;  
Best Local Similarity 40.4%; Pred. No. 2.4e-15;  
Matches 38; Conservative 23; Mismatches 30; Indels 3; Gaps 2;  
QY 22 KDSLQWITDQYNKESDDKYHFRIPRLVKVQVQVTDHLEHLNVEMQWTTCKPPTTN--C 79  
DB 46 KQCLWFAMQYNEKSEDKYFVLVVKTLQALQVNTNLEVLIDVEIARSDCRKPLSTNEIC 105  
QY 80 VPQER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 112  
DB 106 AIQENSKLKRKLSCLSGFLVGALPWNGEFTWMEKCC 139

RESULT 5  
US-09-431-480-2  
; Sequence 2, Application US/09431480  
; Patent No. 6235708  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72  
; CURRENT APPLICATION NUMBER: US/09/431,480  
; CURRENT FILING DATE: 1999-11-01  
; EARLIER APPLICATION NUMBER: 60/109,217  
; EARLIER FILING DATE: 1998-11-20  
; EARLIER APPLICATION NUMBER: 60/156,382  
; EARLIER FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-431-480-2

Query Match 29.8%; Score 187.5; DB 3; Length 141;  
Best Local Similarity 39.6%; Pred. No. 3.1e-15;  
Matches 38; Conservative 20; Mismatches 35; Indels 3; Gaps 2;  
QY 20 YAKDSLQWITDQYNKESDDKYHFRIPRLVKVQVQVTDHLEHLNVEMQWTTCKP--PETT 77  
DB 43 YVQHALWYAMKEYNKASNDLYNFRVVDILKSQEQITDSLEYLYLEVINIARTWCKKIADNE 102  
QY 78 NCV-PQERELHKQVNCFFSVFVFPWFPEQYKILNKSC 112  
DB 103 NCLFQDDPKMKWFCIFIVSSKPKFKLKMALKKC 138

RESULT 6  
US-09-617-302-2  
; Sequence 2, Application US/09617302  
; Patent No. 6245529  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72 C1  
; CURRENT APPLICATION NUMBER: US/09/617,302  
; CURRENT FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/431,480

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 29,4345 Seconds  
(without alignments)  
201.701 Million cell updates/sec

Title: US-09-941-314-3  
Perfect score: 629  
Sequence: 1 YQARKKTFSLVHEWMAVENY.....VFAPWFEQYKILNKSCSSD 115

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214	34.0	142	3	US-09-431-480-3
2	214	34.0	142	3	US-09-617-302-3
3	188.5	30.0	142	3	US-09-431-480-4
4	188.5	30.0	142	3	US-09-617-302-4
5	187.5	29.8	141	3	US-09-431-480-2
6	187.5	29.8	141	3	US-09-617-302-2
7	171.5	27.3	127	4	US-08-849-303-19
8	170.5	27.1	116	4	US-09-775-932-16
9	170.5	27.1	139	2	US-08-791-522-4
10	170.5	27.1	139	3	US-09-314-777-4
11	170.5	27.1	139	4	US-08-849-303-15
12	162	25.8	146	6	5432264-6
13	160.5	25.5	121	4	US-09-775-932-8
14	160.5	25.5	141	3	US-09-886-319A-46
15	160.5	25.5	141	3	US-08-744-138-6
16	160.5	25.5	141	4	US-09-241-376-6
17	160.5	25.5	141	4	US-09-940-497-6
18	160.5	25.5	141	4	US-08-849-303-24
19	156.5	24.9	140	3	US-09-431-480-5
20	156.5	24.9	140	3	US-09-617-302-5
21	156.5	24.9	140	4	US-09-886-319A-48
22	156.5	24.9	140	4	US-08-849-303-18
23	155	24.6	120	4	US-09-775-932-2
24	155	24.6	120	6	5432264-4
25	155	24.6	145	2	US-08-832-535-11
26	155	24.6	146	2	US-08-791-522-3
27	155	24.6	146	3	US-08-744-138-3

28	155	24.6	146	3	US-09-019-485-4	Sequence 4, Appli
29	155	24.6	146	3	US-09-314-777-3	Sequence 3, Appli
30	155	24.6	146	3	US-09-431-480-6	Sequence 6, Appli
31	155	24.6	146	3	US-09-617-302-6	Sequence 6, Appli
32	155	24.6	146	4	US-09-241-376-3	Sequence 3, Appli
33	155	24.6	146	4	US-09-528-436B-3	Sequence 3, Appli
34	155	24.6	146	4	US-09-886-319A-47	Sequence 47, Appli
35	155	24.6	146	4	US-09-940-497-3	Sequence 3, Appli
36	155	24.6	146	4	US-09-976-594-37	Sequence 37, Appli
37	155	24.6	146	4	US-08-849-303-17	Sequence 17, Appli
38	155	24.6	146	5	PCT-US95-07135-9	Sequence 9, Appli
39	153.5	24.4	112	4	US-08-849-303-16	Sequence 16, Appli
40	153.5	24.4	118	4	US-09-775-932-24	Sequence 24, Appli
41	146.5	23.3	121	4	US-09-775-932-4	Sequence 4, Appli
42	146.5	23.3	141	3	US-08-744-138-5	Sequence 5, Appli
43	146.5	23.3	141	3	US-09-431-480-11	Sequence 11, Appli
44	146.5	23.3	141	3	US-09-617-302-11	Sequence 11, Appli
45	146.5	23.3	141	4	US-09-241-376-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-09-431-480-3  
; Sequence 3, Application US/09431480  
; Patent No. 6235708  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72  
; CURRENT APPLICATION NUMBER: US/09/431,480  
; CURRENT FILING DATE: 1999-11-01  
; EARLIER APPLICATION NUMBER: 60/109,217  
; EARLIER FILING DATE: 1998-11-20  
; EARLIER APPLICATION NUMBER: 60/156,382  
; EARLIER FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-431-480-3

Query Match	34.0%	Score 214;	DB 3;	Length 142;
Best Local Similarity	37.1%;	Pred. No. 1.8e-18;		
Matches	43;	Conservative	29;	Mismatches 30; Indels 14; Gaps 4;
Qy	10	SVHEWMAVENY-----AKDSLOWITDQNKESDDKXHFRIKVLKVRQVTDHLE	59	
Db	25	SKNEVKA-QNYFGSINISNANVKQCVFAMKEYNKESDKYFLVDKILHAKLQITDRME	83	
Qy	60	YHLNVEMQWTTCKP--ETTNCVPQER-ELHKVNCFFSVFAVPWPEQYKILNKSC	112	
Db	84	YQIDVOISRNCKKPLNNTENCIPKKPELEKKNKSCFLVGALPWFNGFNLLSKEC	139	

RESULT 2  
US-09-617-302-3  
; Sequence 3, Application US/09617302  
; Patent No. 6245529  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; APPLICANT: Feldhaus, Andrew  
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
; FILE REFERENCE: 98-72 C1  
; CURRENT APPLICATION NUMBER: US/09/617,302  
; CURRENT FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/431,480  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: 60/109,217

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-488899/53.  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts.  
 XX  
 PS Claim 15; SEQ ID NO 21294; 530pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21533-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease.  
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 50 AA;  
 Query Match 41.7%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 4 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVQRQ 53  
 DB 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVQRQ 50  
 RESULT 15  
 AAM67252  
 ID AAM67252 standard; protein; 50 AA.  
 AC AAM67252;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27558.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US0000668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234587P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-488900/53.  
 XX  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.  
 XX  
 PS Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.  
 XX

CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention  
 XX  
 SQ Sequence 50 AA;  
 Query Match 41.7%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 4 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVQRQ 53  
 DB 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVQRQ 50  
 Search completed: March 18, 2004, 14:14:56  
 Job time : 106.276 secs

XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-496933/54.  
 DR  
 XX New spatially-addressable set of single exon nucleic acid probes, useful  
 PT for measuring gene expression in sample derived from human breast,  
 PT comprises number of single exon nucleic acid probes.  
 XX  
 PS Claim 27; SEQ ID NO 15357; 327pp + Sequence Listing; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting the  
 CC probes with a collection of detectably labelled nucleic acids derived  
 CC from mRNA of human breast, and then measuring the label bound to each  
 CC probe of the microarray. The probes are useful for verifying the  
 CC expression of regions of genomic DNA predicted to encode proteins. They  
 CC are useful for gene discovery, and for determining predisposition and/or  
 CC assessing the toxicity of chemical agents on cells. The microarray of  
 CC this invention presents a far greater diversity of probes for measuring  
 CC gene expression, with far less bias than expressed sequence tag  
 CC microarrays. The method is suitable for rapid production of functional  
 CC information from genomic sequence. The present sequence is a peptide  
 CC encoded by a single exon nucleic acid probe of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 50 AA;  
 Query Match 41.7%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 RKKTFLSVHVMVAENVAKDSLQWITDQYNKESDDKYHFRFVLKVRQ 53  
 Db 1 RKKTFLSVHVMVAENVAKDSLQWITDQYNKESDDKYHFRFVLKVRQ 50  
 RESULT 13  
 ABB28913  
 ID ABB28913 standard; peptide; 50 AA.  
 XX  
 AC ABB28913;  
 DT 01-FEB-2002 (first entry)  
 DE Peptide #1564 encoded by breast cell single exon nucleic acid probe.  
 KW Human; microarray; single exon probe; gene expression; breast; disease;  
 XX cancer.  
 XX Homo sapiens.  
 OS  
 XX WO200157271-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US0000662.  
 PF  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-496933/54.  
 DR  
 XX New spatially-addressable set of single exon nucleic acid probes, useful  
 PT for measuring gene expression in sample derived from human breast,  
 PT comprises number of single exon nucleic acid probes.  
 XX  
 PS Claim 27; SEQ ID NO 11881; 327pp + Sequence Listing; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting the  
 CC probes with a collection of detectably labelled nucleic acids derived  
 CC from mRNA of human breast, and then measuring the label bound to each  
 CC probe of the microarray. The probes are useful for verifying the  
 CC expression of regions of genomic DNA predicted to encode proteins. They  
 CC are useful for gene discovery, and for determining predisposition and/or  
 CC assessing the toxicity of chemical agents on cells. The microarray of  
 CC this invention presents a far greater diversity of probes for measuring  
 CC gene expression, with far less bias than expressed sequence tag  
 CC microarrays. The method is suitable for rapid production of functional  
 CC information from genomic sequence. The present sequence is a peptide  
 CC encoded by a single exon nucleic acid probe of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 50 AA;  
 Query Match 41.7%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 RKKTFLSVHVMVAENVAKDSLQWITDQYNKESDDKYHFRFVLKVRQ 53  
 Db 1 RKKTFLSVHVMVAENVAKDSLQWITDQYNKESDDKYHFRFVLKVRQ 50  
 RESULT 14  
 ABB19524  
 ID ABB19524 standard; protein; 50 AA.  
 XX  
 AC ABB19524;  
 DT 23-JAN-2002 (first entry)  
 DE Protein #1523 encoded by probe for measuring heart cell gene expression.  
 DE Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200157274-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US0000666.  
 PF  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA

```
RESULT 10
ABB34086
ID ABB34086 standard; peptide; 50 AA.
XX
XX ABB34086;
AC
XX
XX 04-FEB-2002 (first entry)
DT
XX
XX Peptide #1592 encoded by human foetal liver single exon probe.
DE
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
KW
XX
XX Homo sapiens.
OS
XX WO200157277-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000669.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
XX Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 50 AA;
SQ
Query Match 41.7%; Score 262; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 RKKTFLSVHEWMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 53
Db 1 RKKTFLSVHEWMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50
RESULT 11
AAM27545
ID AAM27545 standard; protein; 50 AA.
XX
XX AAM27545;
AC
XX
XX 17-OCT-2001 (first entry)
DT
XX
XX Peptide #1582 encoded by probe for measuring placental gene expression.
DE
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
```

```
OS Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
XX Claim 27; SEQ ID NO 27814; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders
XX
XX Sequence 50 AA;
SQ
Query Match 41.7%; Score 262; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 RKKTFLSVHEWMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 53
Db 1 RKKTFLSVHEWMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50
RESULT 12
ABB32389
ID ABB32389 standard; peptide; 50 AA.
XX
XX ABB32389;
AC
XX
XX 01-FEB-2002 (first entry)
DT
XX
XX Peptide #5040 encoded by breast cell single exon nucleic acid probe.
DE
XX
XX Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer.
XX
XX Homo sapiens.
OS
XX
XX WO200157271-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000662.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
```

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX SQ Sequence 48 AA;  
 Query Match 43.4%; Score 273; DB 5; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-23;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 NVEMQWTTCKPPTNCVQPERELHKQVNCFFSVFAVPWFQYKILNK 110  
 Db 1 NVEMQWTTCKPPTNCVQPERELHKQVNCFFSVFAVPWFQYKILNK 48

RESULT 8  
 AAU79863  
 ID AAU79863 standard; peptide: 49 AA.  
 XX  
 AC AAU79863;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #11.  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0320230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PS Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 97-98; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX SQ Sequence 49 AA;  
 Query Match 43.2%; Score 272; DB 5; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-23;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 DOYNKESDDKYHFRIFRVLKVQVQVTDHLEVHLNVEMQWTTCKPPTN 78  
 Db 1 DOYNKESDDKYHFRIFRVLKVQVQVTDHLEVHLNVEMQWTTCKPPTN 49

RESULT 9  
 AAM15096  
 ID AAM15096 standard; protein: 50 AA.  
 XX  
 AC AAM15096;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Peptide #1530 encoded by probe for measuring cervical gene expression.  
 XX  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.  
 XX  
 PS Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human cervical epithelial cells.  
 XX  
 PS Claim 27; SEQ ID NO 19922; 487pp; English.  
 XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENPs: see AAI10068-AAI128459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 50 AA;  
 Query Match 41.7%; Score 262; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVQV 53  
 Db 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVQV 50

CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
SQ Sequence 59 AA;

Query Match 53.1%; Score 334; DB 5; Length 59;  
Best Local Similarity 100.0%; Pred. No. 2.3e-30;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 RQVTDHLEYHLNEMQWTTCKPPTNCVPOERLHKQVNCFFSVFAPWPFQYKILNK 110  
Db 1 RQVTDHLEYHLNEMQWTTCKPPTNCVPOERLHKQVNCFFSVFAPWPFQYKILNK 59

RESULT 6  
AAU79864  
ID AAU79864 standard; peptide; 52 AA.  
AC AAU79864;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human cystatin-8 (Zcys8) antigenic fragment #12.  
XX  
DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic peptide.  
XX  
OS Homo sapiens.  
XX  
PN WO200220567-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US026868.  
XX  
PR 01-SEP-2000; 2000US-0230230P.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Holloway JL, Gao Z, Bishop PD;  
XX  
DR WPI; 2002-383044/41.  
XX  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX  
PS Claim 2; Page 98; 100pp; English.  
XX  
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
SQ Sequence 59 AA;

Query Match 53.1%; Score 334; DB 5; Length 59;  
Best Local Similarity 100.0%; Pred. No. 2.3e-30;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 RQVTDHLEYHLNEMQWTTCKPPTNCVPOERLHKQVNCFFSVFAPWPFQYKILNK 110  
Db 1 RQVTDHLEYHLNEMQWTTCKPPTNCVPOERLHKQVNCFFSVFAPWPFQYKILNK 59

RESULT 6  
AAU79864  
ID AAU79864 standard; peptide; 52 AA.  
AC AAU79864;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human cystatin-8 (Zcys8) antigenic fragment #12.  
XX  
DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic peptide.  
XX  
OS Homo sapiens.  
XX  
PN WO200220567-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US026868.  
XX  
PR 01-SEP-2000; 2000US-0230230P.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Holloway JL, Gao Z, Bishop PD;  
XX  
DR WPI; 2002-383044/41.  
XX  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX  
PS Claim 2; Page 98; 100pp; English.  
XX  
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
SQ Sequence 52 AA;

Query Match 45.8%; Score 288; DB 5; Length 52;  
Best Local Similarity 100.0%; Pred. No. 3.4e-25;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 KESDDKXHFRIFRVLKQVRQVTDHLEYHLNEMQWTTCKPPTNCVPOERE 85  
Db 1 KESDDKXHFRIFRVLKQVRQVTDHLEYHLNEMQWTTCKPPTNCVPOERE 52

RESULT 7  
AAU79867  
ID AAU79867 standard; peptide; 48 AA.  
AC AAU79867;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human cystatin-8 (Zcys8) antigenic fragment #15.  
XX  
DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic peptide.  
XX  
OS Homo sapiens.  
XX  
PN WO200220567-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US026868.  
XX  
PR 01-SEP-2000; 2000US-0230230P.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Holloway JL, Gao Z, Bishop PD;  
XX  
DR WPI; 2002-383044/41.  
XX  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX  
PS Claim 2; Page 99; 100pp; English.  
XX  
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's

CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)  
 XX  
 SQ Sequence 137 AA;

Query Match 100.0%; Score 629; DB 5; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-63;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRVLKQVQROVTDHLEY 60  
 DB 23 YQARKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRVLKQVQROVTDHLEY 82  
 QY 61 HLNVMQWTTTCQKPEPTTNCVQRELHKQVNCFFSVFAVPWFQYKILNKSDD 115  
 DB 83 HLNVMQWTTTCQKPEPTTNCVQRELHKQVNCFFSVFAVPWFQYKILNKSDD 137

RESULT 4  
 AAU79865  
 ID AAU79865 standard; peptide; 80 AA.  
 AC AAU79865;  
 XX  
 DT 15-JUL-2002 (first entry)  
 DE Human cystatin-8 (Zcys8) antigenic fragment #13.  
 DE  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 98; 100pp; English.

XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in

CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX  
 SQ Sequence 80 AA;

Query Match 70.9%; Score 446; DB 5; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-43;  
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 QYNKESDDKYHFRVLKQVQROVTDHLEYHLNVMQWTTTCQKPEPTTNCVQRELHKQV 90  
 DB 1 QYNKESDDKYHFRVLKQVQROVTDHLEYHLNVMQWTTTCQKPEPTTNCVQRELHKQV 60  
 QY 91 NCFFSVFAVPWFQYKILNK 110  
 DB 61 NCFFSVFAVPWFQYKILNK 80

RESULT 5  
 AAU79866  
 ID AAU79866 standard; peptide; 59 AA.  
 AC AAU79866;  
 XX  
 DT 15-JUL-2002 (first entry)  
 DE Human cystatin-8 (Zcys8) antigenic fragment #14.  
 DE  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 99; 100pp; English.

XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 CC  
 XX  
 SQ Sequence 115 AA;  
 Query Match 100.0%; Score 629; DB 5; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-63;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 60  
 DB 1 YQARKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 60  
 QY 61 HLNVEQWTTTCQKPEPTTNCVQERLHKQVNCFFSVFVAVPWFQYKILNKSCSSD 115  
 DB 61 HLNVEQWTTTCQKPEPTTNCVQERLHKQVNCFFSVFVAVPWFQYKILNKSCSSD 115

RESULT 2  
 AAU79854  
 ID AAU79854 standard; protein; 117 AA.  
 XX  
 AC AAU79854;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #2.  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic fragment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 94-95; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 CC  
 XX  
 SQ Sequence 117 AA;  
 Query Match 100.0%; Score 629; DB 5; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-63;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 60  
 DB 3 YQARKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 62  
 QY 61 HLNVEQWTTTCQKPEPTTNCVQERLHKQVNCFFSVFVAVPWFQYKILNKSCSSD 115  
 DB 63 HLNVEQWTTTCQKPEPTTNCVQERLHKQVNCFFSVFVAVPWFQYKILNKSCSSD 117

RESULT 3  
 AAU79852  
 ID AAU79852 standard; protein; 137 AA.  
 XX  
 AC AAU79852;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8).  
 XX  
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 DR N-PSDB; ABK49522.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 93-94; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 104.276 Seconds  
(without alignments)  
311.606 Million cell updates/sec

Title: US-09-941-314-3  
Perfect score: 629  
Sequence: 1 YQARKKTFLLSVHEVMAVNY.....VFAVPWFQYKILNKSCSSD 115

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:.\*  
1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001s:.\*  
5: Geneseq2002s:.\*  
6: Geneseq2003as:.\*  
7: Geneseq2003bs:.\*  
8: Geneseq2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	100.0	115	5 AAU79853	Aau79853 Human cys
2	629	100.0	117	5 AAU79854	Aau79854 Human cys
3	629	100.0	137	5 AAU79852	Aau79852 Human cys
4	446	70.9	80	5 AAU79865	Aau79865 Human cys
5	334	53.1	59	5 AAU79866	Aau79866 Human cys
6	288	45.8	52	5 AAU79864	Aau79864 Human cys
7	273	43.4	48	5 AAU79867	Aau79867 Human cys
8	272	43.2	49	5 AAU79863	Aau79863 Human cys
9	262	41.7	50	4 AAM15096	Aam15096 Peptide #
10	262	41.7	50	4 ABB34086	Abb34086 Peptide #
11	262	41.7	50	4 AAM27545	Aam27545 Peptide #
12	262	41.7	50	4 ABB32389	Abb32389 Peptide #
13	262	41.7	50	4 ABB28913	Abb28913 Peptide #
14	262	41.7	50	4 ABB19524	Abb19524 Protein #
15	262	41.7	50	4 AAM67252	Aam67252 Human bon
16	262	41.7	50	4 AAM54871	Aam54871 Human bra
17	262	41.7	50	4 ABG48915	Abg48915 Human liv
18	262	41.7	50	4 AAM02833	Aam02833 Peptide #
19	262	41.7	50	5 ABG36903	Abg36903 Human pep
20	254	40.4	46	5 AAU79860	Aau79860 Human cys
21	214	34.0	142	4 AAE02404	Aae02404 Murine cy
22	214	34.0	142	4 AAE04433	Aae04433 Mouse spe
23	214	34.0	143	6 ADA14374	Ada14374 Mouse spe
24	199	31.6	142	7 ADD46708	Add46708 Rat Prote
25	199	31.6	142	7 ADD46704	Add46704 Rat Prote

ALIGNMENTS

RESULT 1  
AAU79853  
ID AAU79853-standard; protein; 115 AA.  
XX  
AC AAU79853;  
XX

DT 15-JUL-2002 (first entry)  
XX

DE Human cystatin-8 (Zcys8) antigenic fragment #1.  
XX

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KW sperm motility; fertilisation; antigenic fragment.  
XX

OS Homo sapiens.  
XX

XX WO200220567-A2.  
XX

PD 14-MAR-2002.  
XX

PF 29-AUG-2001; 2001WO-US026868.  
XX

PR 01-SEP-2000; 2000US-0230230P.  
XX

XX (ZYMO ) ZYMOGENETICS INC.  
XX

PI Holloway JL, Gao Z, Bishop PD;  
XX

DR WPI; 2002-383044/41.  
XX

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX

XX Claim 2; Page 94; 100pp; English.  
XX

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 3.21429 Seconds  
(without alignments)  
437.389 Million cell updates/sec

Title: US-09-941-314-9

Perfect score: 143  
Sequence: 1 DQYNKSDKXKFRFRVLEKVRQVTD 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues  
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	137	CS11_HUMAN	Q9H112 homo sapien
2	96	67.1	139	CS11_MOUSE	Q9D269 mus musculu
3	77	53.8	142	CS18_HUMAN	O60676 homo sapien
4	76	53.1	142	CS18_MOUSE	P32766 mus musculu
5	70	49.0	148	CYT8_RABIT	O97862 cyctolegus
6	66	46.2	139	CYT8_CHICK	P01038 gallus galli
7	66	46.2	141	CYT8_HUMAN	P09228 homo sapien
8	65	45.5	146	CYT8_MOUSE	P01034 mus musculu
9	65	45.5	146	CYT8_MACMU	O19092 macaca mula
10	65	45.5	146	CYT8_SAISC	O19093 salmistr sci
11	64	44.8	146	CYT8_COTUA	P81061 coturnix co
12	63	44.1	142	CS18_RAT	O88966 rattus norv
13	62	43.4	127	CYT8_RAT	P14841 rattus norv
14	62	43.4	140	CYT8_MOUSE	P14660 mus musculu
15	62	43.4	148	CYT8_BOVIN	P01035 bos taurus
16	61	42.7	129	CYT8_CYPCA	P35481 cyprinus ca
17	56	39.2	162	CYT8_ONCVO	P22085 onchocerca
18	54	37.8	141	CYT8_HUMAN	P01037 homo sapien
19	54	37.8	141	CYT8_HUMAN	P01037 homo sapien
20	53	37.1	141	CYT8_RAT	P08933 rattus norv
21	51	35.7	111	CYT8_BITAR	P08933 rattus norv
22	51	35.7	165	CS18_HUMAN	Q9H114 homo sapien
23	49	34.3	337	CATV_NPVWC	O64181 mamestra co
24	49	34.3	525	HRG_HUMAN	P04186 homo sapien
25	49	34.3	1146	YHC3_YEAST	O91889 spodoptera
26	48	33.6	337	CATV_NPVSE	P04186 homo sapien
27	48	33.6	806	SUS1_ARATH	P10034 arabidopsis
28	48	33.6	917	NIR2_ARATH	P10034 arabidopsis
29	47.5	33.2	505	VP5_AHSV4	Q02166 african hor
30	47	32.9	149	CYT8_HUMAN	P15828 homo sapien
31	47	32.9	224	HLJ1_YEAST	P48353 saccharomyc
32	47	32.9	602	PGH1_MOUSE	P22437 mus musculu
33	47	32.9	602	PGH1_RAT	Q63921 rattus norv

34	46	32.2	734	1	KK08_YEAST	P36004 saccharomyc
35	45.5	31.8	430	1	ENO_BACSU	P37869 bacillus su
36	45	31.5	73	1	EP04_NPYLD	P30326 lymantria d
37	45	31.5	130	1	CYT8_ONCKE	O98967 oncorhynch
38	45	31.5	130	1	CYT8_ONCKE	O91195 oncorhynch
39	45	31.5	137	1	CS19_MOUSE	P52469 mus musculu
40	45	31.5	258	1	UL79_HSV6U	P49034 human herpe
41	45	31.5	803	1	SUS1_ALINGL	P49034 alnus gluti
42	45	31.5	805	1	SUS1_TULGE	O41608 tulipa gesn
43	45	31.5	918	1	NIA_CUCMA	P17569 cucurbita m
44	44.5	31.1	859	1	PMS2_MOUSE	P54279 mus musculu
45	44	30.8	132	1	MERR_BACCE	P22853 bacillus ce

## ALIGNMENTS

RESULT 1  
ID CS11\_HUMAN STANDARD; PRT; 137 AA.  
AC Q9H112; Q9H113;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin II precursor.  
GN CS11 OR CS18L.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_Taxid=9606;  
RN 1  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Garder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dunn P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Levaeslho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McComachie L.J., McKay K., McMurtry A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
-1- SUBCELLULAR LOCATION: Secreted (Potential).  
-1- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Name=1;  
IsoId=Q9H112-1; Sequence=Displayed;  
Name=2;  
IsoId=Q9H112-2; Sequence=VSP\_001260;  
Note=No experimental confirmation available;  
-1- SIMILARITY: Belongs to the cystatin family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AL096677; CAC13170.1; -  
 DR EMBL; AL096677; CAC17423.1; -  
 DR HSSP; P01038; 1A90.  
 DR Genew; HGNC:15959; CST11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF000031; cystatin; 1.  
 DR SMART; SM00043; Cy; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
 DR Thiol protease inhibitor; Signal; Alternative splicing.  
 KW SIGNAL 1 25  
 FT CHAIN 26 137  
 FT SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 93 101 BY SIMILARITY.  
 FT CARBOHYD 131 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARSPLIC 76 110 Missing (in isoform 2).  
 FT /FTID=VSP\_001260.  
 SQ SEQUENCE 137 AA; 16375 MM; CS656C8C3A585C3B CRC64;

Query Match 100.0%; Score 143; DB 1; Length 137;  
 Best local Similarity 100.0%; Pred. No. 1.1e-14;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DQYKESDDKYHFRIRVAVKQROVTD 27  
 Db 52 DQYKESDDKYHFRIRVAVKQROVTD 78

RESULT 2  
 CS11\_MOUSE  
 ID CS11\_MOUSE STANDARD; PRT; 139 AA.  
 AC Q9D269;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin 11 precursor.  
 GN CST11.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kakuoka T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., deBonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Brownstein S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AK020300; BAB32061.1; -  
 DR HSSP; P01034; 1G96.  
 DR MGD; MGI:1925490; Cst11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF000031; cystatin; 1.  
 DR SMART; SM00043; Cy; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
 DR Thiol protease inhibitor; Signal.  
 KW SIGNAL 1 28  
 FT CHAIN 29 139  
 FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 94 102 BY SIMILARITY.  
 FT CARBOHYD 134 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARSPLIC 134 134 Missing (in isoform 2).  
 FT /FTID=VSP\_001260.  
 SQ SEQUENCE 139 AA; 16217 MM; F22BD9815FA32640 CRC64;

Query Match 67.1%; Score 96; DB 1; Length 139;  
 Best local Similarity 61.5%; Pred. No. 1.5e-07;  
 Matches 16; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

OY 1 DQYKESDDKYHFRIRVAVKQROVTD 26  
 Db 53 DEYKESDDLYNFRILIKIMKROVT 78

RESULT 3  
 CST8\_HUMAN  
 ID CST8\_HUMAN STANDARD; PRT; 142 AA.  
 AC O60676;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin  
 DE 8).  
 GN CST8 OR CRES.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=95344753; PubMed=7619504;  
 RA Cornwall G.A., Hann S.R.;  
 RT "Transient appearance of CRES protein during spermatogenesis and  
 RT caput epididymal sperm maturation.";  
 RL Mol. Reprod. Dev. 41:37-46(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones W., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blake S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cooley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dham P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A.M., French L., Garner P.,  
 RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leharasliho M.H., Leverisha M.A., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McComachie L.J., McInay K., McMurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramway H.,  
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Suleton J.B.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epithelium.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF059244; AAC14707.1; -;  
 CC EMBL: AL109954; CAB64234.1; -;  
 CC HSSP: P01034; 1G96.  
 CC Genew: HSCN2480; CST8.  
 CC GO: GO:0004869; P:cysteine protease inhibitor activity; TAS.  
 CC InterPro: IPR000010; Cystatin.  
 CC Pfam: PF00031; cystatin; 1.  
 CC SMART: SM00043; Cy; 1.  
 CC K101 protease inhibitor; Signal; Polymorphism.  
 CC FT SIGNAL 1 21 POTENTIAL.  
 CC FT CHAIN 22 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 CC FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).  
 CC FT DISULFID 95 105 BY SIMILARITY.  
 CC FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT VARIANT 142 142 A -> P (in dbSNP:1054633).  
 CC FT SIGNAL 1 19 /FTID=VAR 014527.  
 CC FT SEQUENCE 142 AA; 16275 MW; 9A512757E0F4E0D CRC64;  
 SO  
 Query Match 53.8%; Score 77; DB 1; Length 142;  
 Best Local Similarity 53.8%; Pred. No. 0.00011;  
 Matches 14; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 Oy 2 QYNKESDCKHFRIFRLVAKVQROYTD 27  
 Db 55 EYNKESDCKVFLVVKTLQALQVNT 80  
 RESULT 4  
 CST8\_MOUSE STANDARD; PRT; 142 AA.  
 ID P32766; O89102;  
 AC P32766; O89102;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-  
 DE related epididymal specific protein) (Cystatin 8).  
 GN CST8 OR CR8.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxId=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CH, and CD-1;  
 RX MEDLINE=9247899; PubMed=10229662;  
 RA Cornwall G.A., Hsie N., Sutton H.G.;  
 RT "Structure, alternative splicing and chromosomal localization of the  
 RT cystatin-related epididymal spermatogenic gene";  
 RL Biochem. J. 340:85-93(1999).  
 RN [2]  
 RP SEQUENCE OF 4-142 FROM N.A.  
 RC TISSUE=Epididymis;  
 RX MEDLINE=93078799; PubMed=1280328;  
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;  
 RT "The CR8 gene: a unique testis-regulated gene related to the cystatin  
 RT family is highly restricted in its expression to the proximal region  
 RT of the mouse epididymis";  
 RL Mol. Endocrinol. 6:1653-1664(1992).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epithelium.  
 CC -1- INDUCTION: Testicular factors or hormones other than androgens  
 CC present in the testicular fluid may be involved in the regulation  
 CC of CR8 gene expression.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF091503; AAC61754.1; -;  
 CC EMBL: AF090691; AAC36316.1; -;  
 CC EMBL: S49926; AAC35390.1; -;  
 CC PIR: A45361; A45361.  
 CC HSSP: P01034; 1G96.  
 CC MGI: MGI:107161; Cat8.  
 CC InterPro: IPR000010; Cystatin.  
 CC Pfam: PF00031; cystatin; 1.  
 CC SMART: SM00043; Cy; 1.  
 CC K101 protease inhibitor; Signal.  
 CC FT SIGNAL 1 19 POTENTIAL.  
 CC FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 CC FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).  
 CC FT DISULFID 95 105 BY SIMILARITY.  
 CC FT CARBOHYD 119 139 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CONFLICT 4 15 PMLSLILFLIP -> GTREOUGESEK (IN REF. 2).  
 CC FT SEQUENCE 142 AA; 16288 MW; 50B446B98F6673E CRC64;  
 SO  
 Query Match 53.1%; Score 76; DB 1; Length 142;  
 Best Local Similarity 50.0%; Pred. No. 0.00016;  
 Matches 13; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
 Oy 2 QYNKESDCKHFRIFRLVAKVQROYTD 27  
 Db 55 EYNKESDCKVFLVVKTLHAKLQITD 80  
 RESULT 5  
 CYTC\_RABIT STANDARD; PRT; 148 AA.  
 ID CYTC\_RABIT  
 AC O97862;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)  
 DE Cystatin C precursor.  
 GN CSTR.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Japanese white; TISSUE=Bone;  
 RX MEDLINE=98424349; PubMed=9753427;  
 RA Kobori M., Ikeda Y., Nara H., Kato M., Kumegawa M., Nojima H.,  
 RA Kawashima H.;  
 RT "Large scale isolation of osteoclast-specific genes by an improved  
 RT method involving the preparation of a subtraced cDNA library.";  
 RL Genes Cells 3:459-475(1998).  
 CC -1- FUNCTION: This is a thiol proteinase inhibitor.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB009342; BAA75921.1; -  
 DR HSSP; P01034; 1696.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
 KM Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT CHAIN 29 148 CYSTATIN C.  
 FT ACT SITE 39 39 REACTIVE SITE.  
 FT SITE 83 87 SECONDARY AREA OF CONTACT.  
 FT DISUFID 101 111 BY SIMILARITY.  
 FT DISUFID 125 145 BY SIMILARITY.  
 SQ SEQUENCE 148 AA; 16346 MW; 1523C83116955B9A CRC64;  
 Query Match 49.0%; Score 70; DB 1; Length 148;  
 Best Local Similarity 45.8%; Pred. No. 0.0014;  
 Matches 11; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 QYNKESDXYFRFRFLKXQROV 25  
 Db 61 EYNNKSNDRYHSRALQVRRARQI 84  
 RESULT 6  
 CYT CHICK  
 ID CYT CHICK STANDARD; PRT; 139 AA.  
 AC P01038;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-OCT-1989 (Rel. 12, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin precursor (Egg-white cystatin).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=90008873; PubMed=2793849;  
 RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;  
 RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,  
 RT and tissue distribution.";  
 RL J. Biol. Chem. 264:17164-17169 (1989).  
 RN [2]  
 RP SEQUENCE OF 24-139.

RX MEDLINE=94178305; PubMed=6712597;  
 RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;  
 RT "Cystatin. Amino acid sequence and possible secondary structure.";  
 RL Biochem. J. 217:813-817 (1984).  
 RN [3]  
 RP SEQUENCE OF 24-139.  
 RX MEDLINE=94110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longner M., Ritonja A., Eropkin M., Borchart U.,  
 RA Machleidt W.;  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 RT of cystatin from chicken egg white.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496 (1983).  
 RN [4]  
 RP CHARACTERIZATION OF PROTEIN.  
 RX MEDLINE=83256421; PubMed=6409085;  
 RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,  
 RA Suter D.C., Barrett A.J.;  
 RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved  
 RT purification from egg white, characterization, and detection in  
 RT chicken serum.";  
 RL Biochem. J. 211:129-138 (1983).  
 RN [5]  
 RP DISULFIDE BONDS.  
 RA Grubb A., Loeffberg H., Barrett A.J.;  
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
 RT cystatin.";  
 RL FEBS Lett. 170:370-374 (1984).  
 RN [6]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=89252033; PubMed=2721673;  
 RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,  
 RA Bode W.;  
 RT "The cysteine proteinase inhibitor chicken cystatin is a  
 RT phosphoprotein.";  
 RL FEBS Lett. 248:162-168 (1989).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=89052676; PubMed=3191914;  
 RA Bode W., Eng R., Musil D., Thiele U., Huber R., Karshkov A.,  
 RA Brzin J., Kos J., Turk V.;  
 RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and  
 RT its possible mode of interaction with cysteine proteinases.";  
 RL EMBO J. 7:2593-2599 (1988).  
 RN [8]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=94087719; PubMed=8263912;  
 RA Dieckmann T., Mitschang L., Hofmann M., Kos J., Turk V.,  
 RA Auerswald E.A., Jeanicke R., Oschkinat H.;  
 RT "The structures of native phosphorylated chicken cystatin and of a  
 RT recombinant unphosphorylated variant in solution.";  
 RL J. Mol. Biol. 234:1048-1059 (1993).  
 CC -1- FUNCTION: This protein binds tightly to and inhibits a variety of  
 CC thiol proteases including ficin, papain, and cathepsins B, C, H,  
 CC and L. Although isolated from egg white, it is also present in  
 CC serum.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J05077; AAA48744.1; -  
 DR PIR; A34456; UDCH.  
 DR PDB; 1CEW; 31-JAN-94.  
 DR PDB; 1A67; 27-MAY-98.  
 DR PDB; 1A90; 17-JUN-98.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.

DR PROSITE; PS00287; CYSTATIN; 1.  
 KW Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 139  
 FT ACT SITE 32 32 CYSTATIN.  
 FT SITE 76 80 REACTIVE SITE.  
 FT DISUFID 94 104 SECONDARY AREA OF CONTACT.  
 FT DISUFID 118 138  
 FT MOD RES 103 103  
 FT STRAND 35 36 PHOSPHORYLATION (PARTIAL).  
 FT TURN 39 40  
 FT HELIX 42 51  
 FT TURN 52 52  
 FT HELIX 53 56  
 FT TURN 57 58  
 FT STRAND 63 77  
 FT STRAND 81 95  
 FT TURN 96 97  
 FT TURN 99 100  
 FT HELIX 101 108  
 FT STRAND 115 125  
 FT TURN 126 129  
 FT STRAND 130 139  
 SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;  
 Query Match: 46.2%; Score 66; DB 1; Length 139;  
 Best Local Similarity: 45.8%; Pred. No. 0.0052;  
 Matches 11; Conservativity 7; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 QYNKESDDKHYFPIFYAKVQROV 25  
 Db 54 EYRASNDKSKRVYVIAKROL 77  
 RESULT 7  
 ID CYTT\_HUMAN STANDARD; PRT; 141 AA.  
 AC P09228; O9UCQ7;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin SA precursor (Cystatin S5).  
 GN Cyst2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88185836; PubMed=3446578;  
 RA Saitoh E., Kim H.-S., Smithies O., Maeda N.;  
 RT "Human cysteine-proteinase inhibitors: nucleotide sequence analysis  
 of three members of the cystatin gene family.";  
 RL Gene 61:329-338(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Cobley S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.U., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Levesaite M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurtry A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Philimore B.J.C.T., Prabhalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,  
 RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [3]  
 RP SEQUENCE OF 21-40.  
 RC TISSUE-Saliva;  
 RX MEDLINE=92138674; PubMed=1778989;  
 RA Isemura S., Saitoh E., Sanada K., Minakata K.;  
 RT "Identification of full-sized forms of salivary (S-type) cystatins  
 (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of  
 cystatin S) in human whole saliva and determination of phosphorylation  
 sites of cystatin S.";  
 RL J. Biochem. 110:648-654(1991).  
 RN [4]  
 RP SEQUENCE OF 25-141.  
 RX MEDLINE=88139220; PubMed=3436950;  
 RA Isemura S., Saitoh E., Sanada K.;  
 RT "Characterization and amino acid sequence of a new acidic cysteine  
 proteinase inhibitor (cystatin SA) structurally closely related to  
 cystatin S, from human whole saliva.";  
 RL J. Biochem. 102:693-704(1987).  
 RN [5]  
 RP PRELIMINARY SEQUENCE OF 25-141.  
 RA Isemura S., Saitoh E., Sanada K., Isemura M., Ito S.;  
 RT "Characterization and amino acid sequence of a new acidic cysteine  
 proteinase inhibitor (cystatin SA) structurally closely related to  
 cystatin S, from human whole saliva.";  
 RL J. Biochem. 102:693-704(1987).  
 RN [6]  
 RP SEQUENCE OF 25-141 FROM N.A.  
 RX MEDLINE=89076505; PubMed=3202964;  
 RA Saitoh E., Isemura S., Sanada K., Kim H.-S., Smithies O., Maeda N.;  
 RT "Cystatin superfamily. Evidence that family II cystatin genes are  
 evolutionarily related to family III cystatin genes.";  
 RL Biol. Chem. Hoppe-Seyler 369:191-197(1988).  
 CC -1- FUNCTION: Thiol protease inhibitor.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M19673; AAA36116.1; -  
 CC EMBL; M19671; AAA36116.1; JOINED.  
 CC EMBL; M19672; AAA36116.1; JOINED.  
 CC EMBL; A1591074; CAC94784.1; -  
 CC PIR; B29632; B29632.  
 CC HSSP; P01034; 1G96.  
 CC GeneW; HGNC:2474; CST2.  
 CC MIM; 123856; -  
 CC GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.  
 CC InterPro; IPR000010; Cystatin.  
 CC Pfam; PF00031; cystatin; 1.  
 CC SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 KW Thiol protease inhibitor; Signal; Multigene family.  
 FT SIGNAL 1 20  
 FT CHAIN 21 141 CYSTATIN SA.

FT ACT\_SITE 32 32 REACTIVE SITE.  
 FT SITE 76 80 SECONDARY AREA OF CONTACT.  
 FT DISUFID 94 104 BY SIMILARITY.  
 FT DISUFID 118 138 BY SIMILARITY.  
 SQ SEQUENCE 141 AA; 16445 MM; EB54915B1B977AA2 CRC64;  
 Query Match 46.2%; Score 66; DB 1; Length 141;  
 Best Local Similarity 41.7%; Pred. No. 0.0052;  
 Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 QYNKESDKYHFRFLKQVROV 25  
 DB 54 EYNKATEDEYRRLRLVLRABEQI 77  
 RESULT 8  
 CYTC\_HUMAN STANDARD; PRT; 146 AA.  
 AC P01034;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)  
 DE (Post-gamma-91obulin).  
 GN CST3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=87219149; PubMed=3495457;  
 RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;  
 RT "Molecular cloning and sequence analysis of cDNA coding for the  
 RT precursor of the human cysteine proteinase inhibitor cystatin C";  
 RL FEBS Lett. 216:229-233 (1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=90303202; PubMed=2363674;  
 RA Abrahamson M., Olafsson I., Palsson A., Ulfbeck M., Lundwall A.,  
 RT "Structure and expression of the human cystatin C gene";  
 RL Biochem. J. 268:287-294 (1990).  
 RN [3]  
 RP SEQUENCE FROM N.A. (HCMVA VARIANT).  
 RC TISSUE=Brain;  
 RX MEDLINE=89235594; PubMed=2541223;  
 RA Levy E., Lopez-Otin C., Ghist J., Galtner D., Frangione B.;  
 RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is  
 RT related to a mutation in the cystatin C gene, an inhibitor of  
 RT cysteine proteases";  
 RL J. Exp. Med. 169:1771-1778 (1989).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89350949; PubMed=2764935;  
 RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.;  
 RA Isemura S., Sanada K.;  
 RT "The human cystatin C gene (CST3) is a member of the cystatin gene  
 RT family which is localized on chromosome 20";  
 RL Biochem. Biophys. Res. Commun. 162:1324-1331 (1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Dickinson D.P., Hewett-Emmett D., Thiesse M.;  
 RT "Acquisition of complex patterns of differential expression in  
 RT epithelial cell populations during the evolution of type 2 cystatin  
 RT genes";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Basley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Codley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahm D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Levesaishi M.H., Leverhna M.A., Lloyd C., Lloyd D.W., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McCormachie L.J., McLeay K., McMurtry A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillips B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showman R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20";  
 RL Nature 414:865-871 (2001).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=2238825; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uebli T.B., Toshitaki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,  
 RA Villalon D.K., Munz D.C., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [8]  
 RP SEQUENCE OF 27-146.  
 RX MEDLINE=8222268; PubMed=6283552;  
 RA Grubb A., Loeffberg H.;  
 RT "Human gamma-trace, a basic microprotein: amino acid sequence and  
 RT presence in the adenylophophysis";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027 (1982).  
 RN [9]  
 RP SEQUENCE OF 27-73.  
 RX MEDLINE=84110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longner M., Ritonja A., Eropkin M., Borchart U.,  
 RA Machleidt W.;  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 RT of cystatin from chicken egg white";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496 (1983).  
 RN [10]  
 RP SEQUENCE OF 27-76.  
 RX MEDLINE=84128015; PubMed=6365094;  
 RA Brzin J., Popovic T., Turk V.;  
 RT "Human cystatin, a new protein inhibitor of cysteine proteinases";  
 RL Biochem. Biophys. Res. Commun. 118:103-109 (1984).  
 RN [11]  
 RP DISULFIDE BONDS.

RA Grubb A., Loeffberg H., Barrett A.J.;  
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
RT cystatin.";  
RL FEBS Lett. 170:370-374(1984).  
RN (12)  
RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.  
RX MEDLINE=21173909; PubMed=11276250;  
RA Janowski R., Kozak M., Jankowska E., Gronka Z., Grubb A.,  
RA Abrahamson M., Jaskolski M.;  
RT "Human cystatin C, an amyloidogenic protein, dimerizes through  
RT three-dimensional domain swapping.";  
RL Nat. Struct. Biol. 8:316-320(2001).  
RN (13)  
RP VARIANT GLN-94.  
RX MEDLINE=23316504; PubMed=1352269;  
RA Abrahamson M., Jonsdottir S., Olafsson I., Jenson O., Grubb A.;  
RT "Hereditary cystatin C amyloid angiopathy: identification of the  
RT disease-causing mutation and specific diagnosis by polymerase chain  
RT reaction based analysis.";  
RL Hum. Genet. 89:377-380(1992).  
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
CC thought to serve an important physiological role as a local  
CC regulator of this enzyme activity.  
CC -1- SUBUNIT: Homodimer.  
CC -1- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,  
CC vas deferens, brain, thymus, and ovary and the lowest in the  
CC submandibular gland.  
CC -1- DISEASE: Defects in CST3 are a cause of hereditary cerebral  
CC hemorrhage with amyloidosis (HCHWA) [MIM:105150]; also known as  
CC cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis  
CC Icelandic type. HCHWA is characterized by a thickening of the  
CC cerebral arteries walls with deposition of material with the  
CC characteristics of amyloid.  
CC -1- SIMILARITY: Belongs to the cystatin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X05607; CAA29096.1; -;  
DR EMBL: X52255; CAA36497.1; -;  
DR EMBL: M27891; AAA52164.1; -;  
DR EMBL: M27889; AAA52164.1; JOINED.  
DR EMBL: M27890; AAA52164.1; JOINED.  
DR EMBL: X61681; CAA43856.2; -;  
DR EMBL: X61682; CAA43856.2; JOINED.  
DR EMBL: X61683; CAA43856.2; JOINED.  
DR EMBL: AF319564; AAK11570.1; -;  
DR EMBL: AL121894; CAC05424.1; -;  
DR EMBL: BC013083; AAI13083.1; -;  
DR PIR: S10216; UDHU.  
DR PDB: 1G96; 06-APR-01.  
DR Genew; HGNC:2475; CST3.  
DR MIM: 604312; -;  
DR MIM: 105150; -;  
DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; cystatin; 1.  
DR SMART: SM00043; CY: 1.  
DR PROSITE: PS00287; CYSTATIN; 1.  
KW Thiol protease inhibitor; Amyloid; Signal; Disease mutation;  
KW Polymorphism; 3D-structure.  
FT SIGNAL 1 26  
FT CHAIN 27 146 CYSTATIN C.  
FT ACT SITE 37 37 REACTIVE SITE.  
FT SITE 81 85 SECONDARY AREA OF CONTACT.  
FT DISULFID 99 109  
FT DISULFID 123 143  
Query Match 45.5%; Score 65; DB 1; Length 146;

Best Local Similarity 41.7%; Pred. No. 0.0077;  
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;  
QY 2 QYNKESDCKYFRIRFLVLYKQROV 25  
DB 59 EYNKASNDWYHSRALQVVRARQOI 82  
RESULT 9  
ID CYTC\_MACMU STANDARD; PRT; 146 AA.  
AC 019092;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin C precursor.  
GN CST3.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP MEDLINE=97054523; PubMed=8898820;  
RX Wei L.H., Walker L.C., Levy E.;  
RA "Cystatin C, Icelandic-like mutation in an animal model of  
RT cerebrovascular beta-amyloidosis.";  
RL Stroke 27:2080-2085(1996).  
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
CC thought to serve an important physiological role as a local  
CC regulator of this enzyme activity.  
CC -1- SIMILARITY: Belongs to the cystatin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U51912; AAB64050.1; -;  
DR HSPB: P01034; 1G96.  
DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; cystatin; 1.  
DR SMART: SM00043; CY: 1.  
DR PROSITE: PS00287; CYSTATIN; 1.  
KW Thiol protease inhibitor; Amyloid; Signal.  
FT SIGNAL 1 26  
FT CHAIN 27 146  
FT ACT SITE 37 37 REACTIVE SITE.  
FT SITE 81 85 SECONDARY AREA OF CONTACT.  
FT DISULFID 99 109 BY SIMILARITY.  
FT DISULFID 123 143 BY SIMILARITY.  
SQ SEQUENCE 146 AA; 15857 MW; F0B3BB774A29D26 CRC64;  
Query Match 45.5%; Score 65; DB 1; Length 146;  
Best Local Similarity 41.7%; Pred. No. 0.0077;  
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;  
QY 2 QYNKESDCKYFRIRFLVLYKQROV 25  
DB 59 EYNKASNDWYHSRALQVVRARQOI 82  
RESULT 10  
ID CYTC\_SAISC STANDARD; PRT; 146 AA.  
AC 019093;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Cystatin C precursor.
GN CST3.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxId=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C, Icelandic-like mutation in an animal model of
RL stroke 27:2080-2085(1996).
CC -1- FUNCTION: As an inhibitor of cysteine proteinase, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CD EMBL; U52028; AAB64051.1; -.
DR HSSP; P01034; 1G96.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 37 146 CYSTATIN C.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 15946 MW; 08196353C0306AA3 CRC64;

Query Match 45.5%; Score 65; DB 1; Length 146;
Best Local Similarity 41.7%; Pred. No. 0.0077;
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIPLVKVOROV 25
DB 59 EYNKASNDKXSRVARIISAKQQL 82

RESULT 11
CYT_COTUA STANDARD; PRT; 116 AA.
AC P81061;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin (Egg-white cystatin).
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxId=93934;
RN [1]
RP SEQUENCE.
RX TISSUE=Egg white; PubMed=9276465;
RA MEDLINE=97420480; PubMed=9276465;
RA GEDHARTZ B., ENGH R.A., MENTELE R., ECKERSKORN C., TORQUATO R.,
RA WITMAN J., KOHL H.J., MACHLEIDT W., FITZ H., AUERWALD B.A.;
RT "Qual cystatin isolation and characterisation of a new member of
RT the cystatin family and its hypothetical interaction with cathepsin
RT B."
RL FEBS Lett. 412:551-558(1997).

```

```

CC -1- FUNCTION: This protein binds tightly to and inhibits papain and
CC cathepsin B.
CC -1- SIMILARITY: Belongs to the cystatin family.
DR HSSP; P01038; 1CEW.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Phosphorylation.
FT ACT_SITE 9 9 REACTIVE SITE.
FT SITE 53 57 SECONDARY AREA OF CONTACT.
FT DISULFID 71 81
FT DISULFID 95 115
FT MOD_RES 80 80 PHOSPHORYLATION.
SQ SEQUENCE 116 AA; 13093 MW; 48248621053A2F70 CRC64;

Query Match 44.8%; Score 64; DB 1; Length 116;
Best Local Similarity 41.7%; Pred. No. 0.0085;
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIPLVKVOROV 25
DB 31 EYNKASNDKXSRVARIISAKQQL 54

RESULT 12
CST8_RAT STANDARD; PRT; 142 AA.
AC O88959;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
DE 8).
GN CST8 OR CR8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Epididymis;
RX MEDLINE=9247899; PubMed=1022962;
RA CORNWALL G.A., HSIA N., SUTTON H.G.;
RT "Structure, alternative splicing and chromosomal localization of the
RT cystatin-related epididymal spermatogenic gene."
RL Biochem. J. 340:85-93(1999).
CC -1- FUNCTION: Performs a specialized role during sperm development and
CC maturation.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CD EMBL; AF090692; AAC6317.1; -.
DR HSSP; P01034; 1G96.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
FT SITE 77 81 PROTEIN.
FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).
FT DISULFID 119 139 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).

```

```

DR HSPB_P01034; 1G96.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART: SM00043; Cy; 1.
DR PROSITE: PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Signal.
FT NON TER 1
FT SIGNAL 1
FT CHAIN <1 7
FT ACT_SITE 8 127 CYSTATIN C.
FT SITE 18 18 REACTIVE SITE.
FT SITE 62 66 SECONDARY AREA OF CONTACT.
FT DISULFID 80 90 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT CONFLICT 25 A -> E (IN REF. 2).
SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;

Query Match 43.4%; Score 62; DB 1; Length 127;
Best Local Similarity 41.7%; Pred. No. 0.019;
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Cy 2 QYNKESDDKTHFRIFRYLVKQROV 25
Db 40 EYNGKSNDAVHSRAIQVRRKOL 63

RESULT 14
CYTC_MOUSE STANDARD; PRT; 140 AA.
ID P21460;
AC P21460;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin C precursor (Cystatin 3).
GN Cg3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=91054523; PubMed=2241983;
RA Solem M., Rawson C., Lindburg K., Barnes D.;
RT "Transforming growth factor beta regulates cystatin C in serum-free
RT mouse embryo (SEME) cells.";
RL Biochem. Biophys. Res. Commun. 172:945-951(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv; TISSUE=Liver;
RX MEDLINE=95137392; PubMed=7835704;
RA Hub C., Nagle J.W., Kozak C.A., Abrahamson M., Karlsson S.;
RT "Structural organization, expression and chromosomal mapping of the
RT mouse cystatin-C-encoding gene (Cstc).";
RL Gene 152:221-226(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ILS, and ISS;
RX MEDLINE=21363810; PubMed=11471062;
RA Ehringer M.A., Thompson J., Contoy O., Xu Y., Yang F., Cammiff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
RT "High-throughput sequence identification of gene coding variants
RT within alcohol-related QTLs.";
RL Mamm. Genome 12:657-663(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Tomilnyk L., Carninci P., Prange C.,

```

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 CC thought to serve an important physiological role as a local  
 CC regulator of this enzyme activity.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: MS9470; AAA63298.1; -;  
 DR EMBL: U10098; ABA1056.1; -;  
 DR EMBL: AF483486; AAL90760.1; -;  
 DR EMBL: AF483487; AAL90761.1; -;  
 DR EMBL: BC002072; AAH02072.1; -;  
 DR PIR: A36163; A36163.  
 DR HSSP: P01034; 1G96.  
 DR MGD: MGI:102519; Cat3.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 DR SMART: SM00043; CY; 1.  
 DR PROSITE: PS00287; CYSTATIN; 1.  
 DR Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 140 CYSTATIN C.  
 FT ACT\_SITE 31 31 REACTIVE SITE.  
 FT SITE 75 79 SECONDARY AREA OF CONTACT.  
 FT DISULFID 93 103 BY SIMILARITY.  
 FT DISULFID 117 137 BY SIMILARITY.  
 FT CONFLICT 16 16 A -> G (IN REF. 1).  
 FT CONFLICT 84 84 L -> F (IN REF. 1).  
 SQ SEQUENCE 140 AA; 15531 MW; 3A563406DD58D0F5 CRC64;  
 QY 2 QYNKESDPKTHFRIFVLKVRQY 25  
 DB 53 EYKNGSNDAYHSPRAIVVPRARKL 76  
 RESULT 15  
 ID CYTC\_BOVIN STANDARD; PRT; 148 AA.  
 AC P01035;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin C precursor (Colostrum thiol proteinase inhibitor).  
 GN CST3.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]

RP SEQUENCE FROM N.A., SEQUENCE OF 66-83, AND CHARACTERIZATION.  
 RC TISSUE=Cerebrospinal fluid, and Choroid plexus;  
 RX MEDLINE=98094199; PubMed=9434110;  
 RA Olsson S.-L., Ek B., Wilm M., Broberg S., Raak L., Björck I.;  
 RT "Molecular cloning and N-terminal analysis of bovine cystatin C  
 RT identification of a full-length N-terminal region."  
 RL Biochem. Biophys. Acta 1343:203-210(1997).  
 RN [2]  
 RP SEQUENCE OF 37-148.  
 RX MEDLINE=85231205; PubMed=3891407;  
 RA Hirado M., Tsunawasa S., Sakiyama F., Nitobe M., Fujii S.;  
 RT "Complete amino acid sequence of bovine colostrum low-Mr cysteine  
 RT proteinase inhibitor."  
 RL FEBS Lett. 186:41-45(1985).  
 CC -1- FUNCTION: This is a thiol proteinase inhibitor.  
 CC -1- MASS SPECTROMETRY: MW=13420; METHOD=MALDI.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Y10811; CAA71771.1; -;  
 DR HSSP: P01034; 1G96.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 DR SMART: SM00043; CY; 1.  
 DR PROSITE: PS00287; CYSTATIN; 1.  
 DR Thiol protease inhibitor; Signal; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 30  
 FT CHAIN 31 148  
 FT MOD\_RES 31 31  
 FT ACT\_SITE 40 40 REACTIVE SITE.  
 FT SITE 84 88 SECONDARY AREA OF CONTACT.  
 FT DISULFID 102 112 BY SIMILARITY.  
 FT DISULFID 126 146 BY SIMILARITY.  
 SQ SEQUENCE 148 AA; 16265 MW; EE740FB37CBB9DCE CRC64;  
 QY 2 QYNKESDPKTHFRIFVLKVRQY 25  
 DB 62 EFNKRSNDAYOSRVVPRARKV 85  
 Search completed: March 18, 2004, 14:16:11  
 Job time : 4.21429 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 16.1786 Seconds

(without alignments)  
526.560 Million cell updates/sec

Title: US-09-941-314-9

Perfect score: 143

Sequence: 1 DQYNKESDDKXHFRIFFVLKVRQVTD 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriopl:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	138	4	Q8WXU6
2	128	89.5	103	4	Q8WXU5
3	92	64.3	139	11	Q8KSA3
4	81	56.6	141	11	Q9DAP1
5	81	56.6	141	11	Q80ZNS
6	66	46.2	112	13	Q98SR4
7	66	46.2	112	13	Q98SR3
8	64	44.8	81	6	Q29212
9	62	43.4	140	11	Q9BPY9
10	57	39.9	148	5	Q9NH55
11	56	39.2	109	5	Q9TY65
12	56	39.2	161	5	Q16159
13	53	37.8	1320	5	Q81LJ3
14	53	37.1	464	13	Q80125
15	51	35.7	428	10	Q9FYQ9
16	50.5	35.3	394	5	Q9T257

17	50	35.0	108	3	Q9P6S2	Q9P6S2 schizosach
18	50	35.0	319	16	Q9CJ05	Q9CJ05 lactococcus
19	50	35.0	393	10	Q9ZMA1	Q9ZMA1 arabidopsis
20	50	35.0	393	10	Q8GZ44	Q8GZ44 arabidopsis
21	45	35.0	425	3	Q12700	Q12700 debaryomyce
22	49	34.3	205	10	Q946X3	Q946X3 prunus pers
23	49	34.3	341	12	Q8JWC4	Q8JWC4 mamestra co
24	49	34.3	403	5	Q9GPR6	Q9GPR6 dictyosteli
25	49	34.3	442	10	Q9SU04	Q9SU04 arabidopsis
26	49	34.3	498	5	Q16454	Q16454 caenorhabdi
27	49	34.3	2646	12	Q9WPZ7	Q9WPZ7 gill-asso
28	48	33.6	91	5	Q9VHP3	Q9VHP3 drosophila
29	48	33.6	109	16	Q7TU55	Q7TU55 prochloroco
30	48	33.6	140	11	Q80Y72	Q80Y72 mus musculu
31	48	33.6	443	5	Q8ICX8	Q8ICX8 plasmodium
32	48	33.6	522	13	Q7ZU78	Q7ZU78 brachydanio
33	48	33.6	664	17	Q8TKV0	Q8TKV0 methanosa
34	48	33.6	804	10	Q9AVR8	Q9AVR8 pisum sativ
35	48	33.6	917	10	Q94C76	Q94C76 arabidopsis
36	48	33.6	2133	5	Q813V5	Q813V5 plasmodium
37	47.5	33.2	133	11	Q9D264	Q9D264 mus musculu
38	47.5	33.2	680	10	Q7XBR4	Q7XBR4 oryza sativ
39	47.5	33.2	868	10	Q7XBR6	Q7XBR6 oryza sativ
40	47	32.9	125	5	Q25620	Q25620 onchocerca
41	47	32.9	370	10	Q9M1J6	Q9M1J6 arabidopsis
42	47	32.9	462	13	Q7ZY91	Q7ZY91 xenopus lae
43	47	32.9	462	13	Q7SYH2	Q7SYH2 xenopus lae
44	47	32.9	465	13	Q801B5	Q801B5 xenopus lae
45	47	32.9	606	6	Q97554	Q97554 oryctolagus

## ALIGNMENTS

RESULT 1	Q8WXU6	PRELIMINARY;	PRT;	138 AA.
ID	Q8WXU6			
AC	Q8WXU5			
DT	01-MAR-2002 (TREMBlrel. 20, Created)			
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)			
DT	01-JUN-2003 (TREMBlrel. 24, Last annotation update)			
DE	SC13.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RA	Hamil K.G., Liu Q., Zhang Y.-L., French P.S., Hall S.H.			
RT	"SC13: A novel epididymal specific member of the cystatin family."			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF35480; AAL71991.1; -			
DR	GO; GO:0004869; F:Cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR00010; Cystatin.			
DR	Pfam; PF00031; Cystatin; 1.			
DR	SMART; SM00043; CY; 1.			
DR	SEQUENCE 138 AA; 16506 MW; B49440ACA3585C64 CRC64;			
SO	SEQUENCE			
Query Match	100.0%; Score 143; DB 4; Length 138;			
Best Local Similarity	100.0%; Pred. No. 4, 4e-14;			
Matches	27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1 DQYNKESDDKXHFRIFFVLKVRQVTD 27			
Db	53 DQYNKESDDKXHFRIFFVLKVRQVTD 79			
RESULT 2	Q8WXU5	PRELIMINARY;	PRT;	103 AA.
ID	Q8WXU5			
AC	Q8WXU5			
DT	01-MAR-2002 (TREMBlrel. 20, Created)			
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)			

```

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE SC13delta.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335461; AAL71992.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;

Query Match 89.5%; Score 128; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 6.5e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFRVLKQVQRTD 24
Db 53 DQYNKESDDKXHFRIFRVLKQVQRTD 76

RESULT 3
08K5A3 PRELIMINARY; PRT; 139 AA.
AC 08K5A3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cystatin 11.
GN CST11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley;
RA Hamil K.G., Hall S.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF501290; AAM21709.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 139 AA; 16686 MW; E1E36DB7864D08C CRC64;

Query Match 64.3%; Score 92; DB 11; Length 139;
Best Local Similarity 51.9%; Pred. No. 3e-06;
Matches 14; Conservative 11; Mismatches 2; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFRVLKQVQRTD 27
Db 53 EYVKKSEDLNFRILRIKIKKQMTN 79

RESULT 4
09DAP1 PRELIMINARY; PRT; 141 AA.
AC 09DAP1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE 1700006C19Rik protein.
GN 1700006C19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai U., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi U., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pasole G., Quackenbush J.,
RA Schriml U.M., Staudl F., Suzuki R., Tomita M., Wagner U., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seye T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK005665; BAB24175.1; -.
DR HSSP; P01038; ICEW.
DR MGD; MGI:1916544; 1700006C19Rik.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 141 AA; 16811 MW; C20FA0D8A1AC378C CRC64;

Query Match 56.6%; Score 81; DB 11; Length 141;
Best Local Similarity 53.8%; Pred. No. 0.00015;
Matches 14; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKQVQRTD 27
Db 54 EYVKKSEDLNFRVLDIKSQEQITD 79

RESULT 5
080ZNS PRELIMINARY; PRT; 141 AA.
AC 080ZNS;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE RIKEN CDNA 1700006C19 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Testicle;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048681; AAH48681.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C/M.
DR Pfam; PF00031; Cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 141 AA; 16825 MW; C20FA0D8A884951F CRC64;

Query Match 56.6%; Score 81; DB 11; Length 141;
Best Local Similarity 53.8%; Pred. No. 0.00015;
Matches 14; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKQVQRTD 27

```

Db 54 EYKASNDLYNFRVNDLKSQEQITD 79

## RESULT 6

Q96SR4 PRELIMINARY; PRT; 112 AA.

AC Q96SR4; 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Cystatin (Fragment).  
 OS Acipenser sinensis (Chinese sturgeon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;  
 OC Acipenser.  
 NCBI\_TaxID=61970;  
 RN NCBI\_TaxID=61970;  
 RP TISSUE=Liver;  
 RC TISSUE=Liver;  
 RA Bai J., Lao H., Ye X., Li Y., Lou J.;  
 RT "Molecular cloning and sequence analysis of cystatin cDNA from two  
 species of sturgeons."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF34610; AAK16731.1; -.  
 DR HSSP; P01038; 1A90.  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 112 AA; 12231 MW; 48CEBFD8A08C00 CRC64;

Query Match 46.2%; Score 66; DB 13; Length 112;  
 Best Local Similarity 50.0%; Pred. No. 0.024;  
 Matches 12; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDDKXTHPRIFRYLKYQROV 25  
 Db 26 EFNKASNDMYIHRVSKVKVQKQV 49

## RESULT 7

Q96SR3 PRELIMINARY; PRT; 112 AA.

AC Q96SR3; 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Cystatin (Fragment).  
 OS Acipenser schrenckii (Amur sturgeon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;  
 OC Acipenser.  
 NCBI\_TaxID=111304;  
 RN NCBI\_TaxID=111304;  
 RP TISSUE=Liver;  
 RC TISSUE=Liver;  
 RA Bai J., Lao H., Ye X., Li Y., Lou J.;  
 RT "Molecular cloning and sequence analysis of cystatin cDNA from two  
 species of sturgeons."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF34611; AAK16732.1; -.  
 DR HSSP; P01038; 1A90.  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 112 AA; 12231 MW; 48CEBFD8A08C00 CRC64;

Query Match 46.2%; Score 66; DB 13; Length 112;

Best Local Similarity 50.0%; Pred. No. 0.024;  
 Matches 12; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDDKXTHPRIFRYLKYQROV 25  
 Db 26 EFNKASNDMYIHRVSKVKVQKQV 49

## RESULT 8

Q29212 PRELIMINARY; PRT; 81 AA.

AC Q29212; 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Cystatin C precursor (Fragment).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 NCBI\_TaxID=9823;  
 RN NCBI\_TaxID=9823;  
 RP TISSUE=Small intestine;  
 RC TISSUE=Small intestine;  
 RA Winteroe A.K., Fredholm M., Davies W.;  
 RT "Evaluation and characterization of a porcine small intestine CDNA  
 library."  
 RL Mamm. Genome 7:509-517 (1996).  
 DR EMBL; F14794; CAA23262.1; -.  
 DR HSSP; P01034; 1G96.  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 FT NON\_TER 81  
 SQ SEQUENCE 81 AA; 8771 MW; 25E83E431135B8A7 CRC64;

Query Match 44.8%; Score 64; DB 6; Length 81;  
 Best Local Similarity 52.2%; Pred. No. 0.035;  
 Matches 12; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 QYNKESDDKXTHPRIFRYLKYQRO 24  
 Db 59 EYKASXVAXHXXRLVLYXVKQ 81

## RESULT 9

Q96PX9 PRELIMINARY; PRT; 140 AA.

AC Q96PX9; 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Cystatin C.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN NCBI\_TaxID=10090;  
 RP TISSUE=Liver;  
 RC STRAIN=BALB/c;  
 RA Taupin P.J., Ray J., Fischer W.H., Suh S.T., Hakanson K., Grubb A.,  
 RT "pGF-2-Responsive neural stem cell proliferation requires CCG, a novel  
 autocrine/paracrine cofactor."  
 RL Neuron 28:385-397 (2000).  
 DR EMBL; AF311741; AAG40283.1; -.  
 DR HSSP; P01034; 1G96.  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.

FT CHAIN 21 140 CYSTATIN C.  
 FT VARIANT 16 A -> G.  
 FT VARIANT 84 L -> F.  
 SQ SEQUENCE 140 AA; 15517 MW; 3A563406DD58D785 CRC64;

Query Match 43.4%; Score 62; DB 11; Length 140;  
 Best Local Similarity 41.7%; Pred. No. 0.12;  
 Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;  
 Oy 2 QYNKESDDKYHFRIRVYLKVQROV 25  
 Db 53 EYKNGSDAYHSRAIQVVRARQOL 76

## RESULT 10

Q9NH95 PRELIMINARY; PRT; 148 AA.  
 AC Q9NH95, (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Is-cystatin.  
 OS Litomosoides sigmodontis.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Litomosoides.  
 OX NCBI\_TaxID=42156;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pfaff A.W., Hoffmann W.H., Taylor D.W., Schulz-Key H.;  
 RT "Characterization and immunological properties of a cysteine protease  
 inhibitor of the filarial parasite Litomosoides sigmodontis."  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF229173; AAF35896.1; -  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 FT CHAIN 25 148 LS-CYSTATIN.  
 SQ SEQUENCE 148 AA; 16686 MW; 2950AA89CA5339C9 CRC64;

Query Match 39.9%; Score 57; DB 5; Length 148;  
 Best Local Similarity 50.0%; Pred. No. 0.77;  
 Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
 Oy 4 NKESDDKYHFRIRVYLKVQROV 25  
 Db 62 NQOSNDAYHLMPIKLVKSSQV 83

## RESULT 11

Q9TY65 PRELIMINARY; PRT; 109 AA.  
 AC Q9TY65, (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Antigen maltese binding protein (Fragment).  
 OS Onchocerca volvulus.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Onchocerca.  
 OX NCBI\_TaxID=6282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94336252; PubMed=8059358;  
 RA Tremblome K.R., Tree T.I., Gillespie A.J., Guderian R., Maizels R.M.,  
 Bradley J.E.;  
 RT "Heterogeneity of Igg antibody responses to cloned Onchocerca volvulus  
 antigens in microfilidermia positive individuals from Esmeraldas  
 Province, Ecuador."  
 RL Parasite Immunol. 16:201-209(1994).  
 DR EMBL; S71364; AAC60509.1; -  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 FT NON\_TPR 1 1  
 SQ SEQUENCE 109 AA; 12701 MW; 3C6B5EF14D8082B8 CRC64;

Query Match 39.2%; Score 56; DB 5; Length 109;  
 Best Local Similarity 45.5%; Pred. No. 0.8;  
 Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 Oy 4 NKESDDKYHFRIRVYLKVQROV 25  
 Db 24 NQOSNDAYHLMPIKLVKSSQV 45

## RESULT 12

O16159 PRELIMINARY; PRT; 161 AA.  
 AC O16159, (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Cystatin-type cysteine proteinase inhibitor.  
 GN BM-CPI-2.  
 OS Brugia malayi (filarial nematode worm).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Brugia.  
 OX NCBI\_TaxID=6279;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gregory W.F., Blaxter M.L., Maizels R.M.;  
 RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gregory W.F., Maizels R.M.;  
 RT "Two distinct cystatin-type cysteine protease inhibitors from the  
 parasitic nematode Brugia malayi."  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF177193; AAB69857.1; -  
 DR EMBL; AF177193; AAB51086.1; -  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 SQ SEQUENCE 161 AA; 18406 MW; 8081351BBE226EB5 CRC64;

Query Match 39.2%; Score 56; DB 5; Length 161;  
 Best Local Similarity 45.5%; Pred. No. 1.2;  
 Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 Oy 4 NKESDDKYHFRIRVYLKVQROV 25  
 Db 73 NQOSNDAYHLMPIKLVKSSQV 94

## RESULT 13

O81LJ3 PRELIMINARY; PRT; 1320 AA.  
 AC O81LJ3, (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN PF14\_0250.  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3D7;  
 RX MEDLINE=22255705; PubMed=12368864;

RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
 RA Carlson J.M., Pain A., Nelson K.B., Bowman S., Paulsen I.T., James K.,  
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,  
 RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Anguoli S.,  
 RA Portier M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,  
 RA Martin D.M.A., Fairhead A.H., Fraumholz M.J., Roos D.S., Ralph S.A.,  
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,  
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
 RA Fraser C.M., Barrall B.,  
 RT "Genome sequence of the human malaria parasite Plasmodium  
 falciparum."  
 RL Nature 419:498-511 (2002).  
 DR EMBL: AEO14820; AAN36863.1; "-"  
 DR GO: GO:0003824; F:catalytic activity; IEA.  
 DR GO: GO:0004806; F:triacylglycerol lipase activity; IEA.  
 DR GO: GO:0006629; F:lipid metabolism; IEA.  
 DR InterPro: IPR002921; Lipase\_3.  
 DR InterPro: IPR008262; Lipase\_AS.  
 DR Pfam: PF01764; Lipase\_3; 1.  
 DR PROSITE: PS00120; LIPASE\_SBR; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1320 AA; 155667 MW; 0F06B631D32DF7AF CRC64;

Query Match 37.8%; Score 54; DB 5; Length 1320;  
 Best Local Similarity 36.0%; Pred. No. 22;  
 Matches 9; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 3 YNKSDDKXHFRIFRVLKVOQVTD 27  
 DB 554 YDKREKQYEIAINMLIKYHKELTD 578

RESULT 14  
 Q80125 PRELIMINARY; PRT; 464 AA.  
 AC 080125;  
 DT 01-JUN-2003 (TRENBLrel. 24, Created)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Fetuin long form.  
 OS Cyrtinus carpio (Common carp).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Cyprinus.  
 OC NCBI\_TaxID=7962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Teal P.-L., Chang G.-D., Huang C.-J.;  
 RT "Purification and cloning of carp fetuin."  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY225965; AAC74862.1; "-"  
 DR GO: GO:0005874; C:microtubule; IEA.  
 DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR GO: GO:0005198; F:structural molecule activity; IEA.  
 DR GO: GO:0007018; P:microtubule-based movement; IEA.  
 DR InterPro: IPR002453; Beta\_tubulin.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; Cystatin; 1.  
 DR SMART: SM00043; CY; 1.  
 DR PROSITE: PS00228; TUBULIN\_B AUTOREG; 1.  
 SQ SEQUENCE 464 AA; 51698 MW; 7A54F71E44050895 CRC64;

Query Match 37.1%; Score 53; DB 13; Length 464;  
 Best Local Similarity 39.1%; Pred. No. 11;  
 Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKVOQVTD 24  
 DB 166 KFNKESDHKSIFKLMVGRISTQ 188

RESULT 15  
 Q9FYQ9

ID Q9FYQ9 PRELIMINARY; PRT; 428 AA.  
 AC Q9FYQ9;  
 DT 01-MAR-2001 (TRENBLrel. 16, Created)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Similarity to helicas.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Columbia;  
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP000421; BAB10021.1; "-"  
 DR GO: GO:0004386; F:helicase activity; IEA.  
 KW Helicase.  
 SQ SEQUENCE 428 AA; 49107 MW; 811FA8B2E790E09 CRC64;

Query Match 35.7%; Score 51; DB 10; Length 428;  
 Best Local Similarity 39.1%; Pred. No. 20;  
 Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKVOQVTD 24  
 DB 89 QYVNHKEKDHILKFLSTLVNQOK 111

Search completed: March 18, 2004, 14:21:28  
 Job time : 17.1786 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 41.7103 Seconds

(without alignments)  
311.606 Million cell updates/sec

Title: US-09-941-314-10

Sequence: 1 NKESDDDKHFRIRFLVKYQR.....EYHLNVEQMWTTCQKPEPTN 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*\n2: geneseqp1990s:\*\n3: geneseqp2000s:\*\n4: geneseqp2001s:\*\n5: geneseqp2002s:\*\n6: geneseqp2003as:\*\n7: geneseqp2003bs:\*\n8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	100.0	46	5	AAU79860 Human cys
2	254	100.0	49	5	AAU79863 Human cys
3	254	100.0	80	5	AAU79865 Human cys
4	254	100.0	115	5	AAU79853 Human cys
5	254	100.0	117	5	AAU79854 Human cys
6	254	100.0	137	5	AAU79852 Human cys
7	254	97.6	52	5	AAU79864 Human cys
8	154	60.6	59	5	AAU79866 Human cys
9	149	58.7	33	5	AAU79862 Human cys
10	138	54.3	24	5	AAU79861 Human cys
11	125	49.2	27	5	AAU79859 Human cys
12	125	49.2	35	5	AAU79858 Human cys
13	118	46.5	92	2	AAW78259 Fragment
14	118	46.5	113	6	ADA57563 Human sec
15	118	46.5	113	6	ADA41457 Human sec
16	118	46.5	113	7	ADC74577 Human sec
17	118	46.5	113	7	ADD38088 Human sec
18	118	46.5	114	2	AAW78153 Human sec
19	118	46.5	123	2	AAW78260 Fragment
20	118	46.5	142	2	AAW78258 Fragment
21	118	46.5	142	4	AAE02405 Human cys
22	118	46.5	142	4	AAE04434 Human cys
23	118	46.5	142	6	ADA57231 Human sec
24	118	46.5	142	6	ADA41112 Human sec
25	118	46.5	142	7	ADC74335 Human sec

26	118	46.5	142	7	ADD37980 Human sec
27	118	46.5	142	7	ADD46706 Human pro
28	118	46.5	142	7	ADD46710 Human pro
29	113	44.5	141	3	AAV96576 Murine cy
30	113	44.5	141	4	AAE02403 Murine cy
31	113	44.5	141	4	AAE04432 Mouse tes
32	111	43.7	142	4	AAE02404 Murine cy
33	111	43.7	142	4	AAE04433 Mouse cys
34	111	43.7	143	6	ADA14374 Mouse spe
35	110	43.3	50	4	AAW15096 Peptide #
36	110	43.3	50	4	ABW34086 Peptide #
37	110	43.3	50	4	AAW27545 Peptide #
38	110	43.3	50	4	ABW32389 Peptide #
39	110	43.3	50	4	ABW32389 Peptide #
40	110	43.3	50	4	ABW19524 Protein #
41	110	43.3	50	4	AAW67252 Human don
42	110	43.3	50	4	AAW54871 Human bra
43	110	43.3	50	4	ABG48915 Human liv
44	110	43.3	50	4	AAW02833 Peptide #
45	110	43.3	50	5	ABG36903 Human pep

## ALIGNMENTS

RESULT 1  
AAU79860 standard; peptide; 46 AA.  
ID AAU79860; AAU79860; (first entry)  
AC 15-UL-2002 (first entry)  
XX 15-UL-2002 (first entry)  
DT Human cystatin-8 (Zcys8) antigenic fragment #8.  
XX Human cystatin-8 (Zcys8) antigenic fragment #8.  
DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
XX sperm motility; fertility; antigenic peptide.  
KW sperm motility; fertility; antigenic peptide.  
XX Homo sapiens.  
OS WO200220567-A2.  
PN 14-MAR-2002.  
PD 29-AUG-2001; 2001WO-US026868.  
XX 01-SEP-2000; 2000US-0230230P.  
PR (ZYMO ) ZYMOGENETICS INC.  
XX Holloway JL, Gao Z, Bishop PD;  
XX WPI; 2002-383044/41.  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX Claim 2; Page 97; 100pp; English.  
XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
XX polypeptide (1). (1) is useful for: inhibiting cancer procoagulant  
XX protein in an individual and thus inhibiting the thrombotic events  
XX associated with cancer; promoting spermatogenesis, modulating seminal  
XX fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
XX motility and fertilisation; and as antigenic peptides for generating  
XX antibodies. Zcys8 is useful as research reagent for characterising sites  
XX of interaction between Zcys8 and its receptor. Zcys8 is useful in  
XX enhancing fertilisation during assisted reproduction in humans and in  
XX animals. Anti-(1) antibodies are useful to screen biological samples like  
XX blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
XX presence of Zcys8. The antibodies are also useful to isolate large

quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes. The polynucleotide encoding (I) is useful to detect and to localise the expression of a Zcys8 gene in a biological sample and Zcys8 oligonucleotide probes are useful for in vivo diagnosis. The polynucleotide encoding (I) is useful in determining whether a subject's chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements and genetic alterations that inactivate the Zcys8 gene. This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

Sequence 46 AA;

Query Match 100.0%; Score 254; DB 5; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.9e-28;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDCKYHFRIFRLVKQROYTDHLEHYHLNVEQMWTTCQKPEPTN 46  
DB 1 NKESDCKYHFRIFRLVKQROYTDHLEHYHLNVEQMWTTCQKPEPTN 46

RESULT 2

AAU79863 standard; peptide, 49 AA.

AAU79863;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #11.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis; spermatogenesis; seminal fluid viscosity; cryopreserved sperm; sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

WO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-38304/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.

Claim 2; Page 97-98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8) polypeptide (I). (I) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcys8 is useful as research reagent for characterising sites of interaction between Zcys8 and its receptor. Zcys8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(I) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcys8. The antibodies are also useful to isolate large quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes. The polynucleotide encoding (I) is useful to detect and to localise the expression of a Zcys8 gene in a biological sample and Zcys8 oligonucleotide probes are useful for in vivo diagnosis. The

polynucleotide encoding (I) is useful in determining whether a subject's chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements and genetic alterations that inactivate the Zcys8 gene. This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

Sequence 49 AA;

Query Match 100.0%; Score 254; DB 5; Length 49;  
Best Local Similarity 100.0%; Pred. No. 2.1e-28;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDCKYHFRIFRLVKQROYTDHLEHYHLNVEQMWTTCQKPEPTN 46  
DB 4 NKESDCKYHFRIFRLVKQROYTDHLEHYHLNVEQMWTTCQKPEPTN 49

RESULT 3

AAU79865 standard; peptide, 80 AA.

AAU79865;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #13.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis; spermatogenesis; seminal fluid viscosity; cryopreserved sperm; sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

WO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-38304/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.

Claim 2; Page 98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8) polypeptide (I). (I) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcys8 is useful as research reagent for characterising sites of interaction between Zcys8 and its receptor. Zcys8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(I) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcys8. The antibodies are also useful to isolate large quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes. The polynucleotide encoding (I) is useful to detect and to localise the expression of a Zcys8 gene in a biological sample and Zcys8 oligonucleotide probes are useful for in vivo diagnosis. The polynucleotide encoding (I) is useful in determining whether a subject's chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX  
 SQ Sequence 80 AA;

Query Match 100.0%; Score 254; DB 5; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-28;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKSDDKTHFRIRVLKVGROVTDHLEHYLNVEMQWTTCKPPTTN 46  
 |||||  
 Db 3 NKSDDKTHFRIRVLKVGROVTDHLEHYLNVEMQWTTCKPPTTN 48

RESULT 4  
 ID AAU79853 standard; protein; 115 AA.

AC AAU79853;

XX 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #1.

KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 XX sperm motility; fertilisation; antigenic fragment.

OS Homo sapiens.

PN WO200220567-A2.

XX 14-MAR-2002.

PF 29-AUG-2001; 2001WO-US026868.

PR 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 94; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)

XX Sequence 115 AA;

Query Match 100.0%; Score 254; DB 5; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-28;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKSDDKTHFRIRVLKVGROVTDHLEHYLNVEMQWTTCKPPTTN 46  
 |||||  
 Db 33 NKSDDKTHFRIRVLKVGROVTDHLEHYLNVEMQWTTCKPPTTN 78

RESULT 5  
 ID AAU79854 standard; protein; 117 AA.

AC AAU79854;

XX 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #2.

KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 XX sperm motility; fertilisation; antigenic fragment.

OS Homo sapiens.

PN WO200220567-A2.

XX 14-MAR-2002.

PF 29-AUG-2001; 2001WO-US026868.

PR 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 94-95; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)

XX Sequence 117 AA;

Query Match 100.0%; Score 254; DB 5; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-28;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIFRVLKVRQVTDHLEHYLNVEMQWTTCCQPEPTN 46  
 |||||  
 DB 35 NKESDDKXHFRIFRVLKVRQVTDHLEHYLNVEMQWTTCCQPEPTN 80

RESULT 6  
 ID AAU79852  
 AAU79852 standard; protein; 137 AA.

AC AAU79852;  
 XX  
 XX 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8).  
 XX  
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200220567-A2.  
 XX  
 XX 14-MAR-2002.  
 XX  
 PD 29-AUG-2001; 2001MO-US026868.  
 XX  
 PF 01-SEP-2000; 2000US-0230230P.  
 XX  
 PR (ZYMO) ZYMOGENETICS INC.  
 PA  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 XX WPI; 2002-383044/41.  
 DR N-PSDB; ABR49522.  
 DR  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 PS  
 PS Claim 2; Page 93-94; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)

SQ Sequence 137 AA;

Query Match 100.0%; Score 254; DB 5; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 7, 1e-28;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIFRVLKVRQVTDHLEHYLNVEMQWTTCCQPEPTN 46

DB 55 NKESDDKXHFRIFRVLKVRQVTDHLEHYLNVEMQWTTCCQPEPTN 100  
 |||||

RESULT 7  
 ID AAU79864  
 AAU79864 standard; peptide; 52 AA.

AC AAU79864;  
 XX  
 XX 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #12.  
 XX  
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KW sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200220567-A2.  
 XX  
 XX 14-MAR-2002.  
 XX  
 PD 29-AUG-2001; 2001MO-US026868.  
 XX  
 PF 01-SEP-2000; 2000US-0230230P.  
 XX  
 PR (ZYMO) ZYMOGENETICS INC.  
 PA  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 XX WPI; 2002-383044/41.  
 DR  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 PS  
 PS Claim 2; Page 98; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

SQ Sequence 52 AA;

Query Match 97.6%; Score 248; DB 5; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 1, 6e-27;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KESDDKXHFRIFRVLKVRQVTDHLEHYLNVEMQWTTCCQPEPTN 46  
 |||||  
 DB 1 KESDDKXHFRIFRVLKVRQVTDHLEHYLNVEMQWTTCCQPEPTN 45





PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 96; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 SO Sequence 35 AA:  
 Query Match 49.2%; Score 125; DB 5; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-10;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NKESDDKYHFRIRVLKVRQVTD 24  
 DB 12 NKESDDKYHFRIRVLKVRQVTD 35  
 ID AAW78259 standard; protein; 92 AA.  
 AC AAW78259;  
 XX  
 DT 13-APR-1999 (first entry)  
 XX  
 DE Fragment of human secreted protein encoded by gene 28.  
 XX  
 KM Human; secreted protein; fusion protein; gene therapy; protein therapy;  
 KM diagnosis; tissue; cancer; neurodegenerative disorder; leukaemia;  
 KM developmental abnormality; foetal deficiency; blood; allergy; renal;  
 KM immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KM inflammation; ischemic shock; Alzheimer's disease; osteoarthritis; AIDS;  
 KM cognitive disorder; schizophrenia; prostate; obesity; osteolact; thymus;  
 KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KM endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 XX  
 OS Homo sapiens.

XX  
 PN WO9856804-A1.  
 XX  
 PD 17-DEC-1998.  
 XX  
 PF 11-JUN-1998; 98WO-US012125.  
 XX  
 PR 13-JUN-1997; 97US-0049547P.  
 PR 13-JUN-1997; 97US-0049548P.  
 PR 13-JUN-1997; 97US-0049549P.  
 PR 13-JUN-1997; 97US-0049550P.  
 PR 13-JUN-1997; 97US-0049551P.  
 PR 13-JUN-1997; 97US-0049552P.  
 PR 13-JUN-1997; 97US-0049553P.  
 PR 13-JUN-1997; 97US-0049554P.  
 PR 13-JUN-1997; 97US-0049555P.  
 PR 13-JUN-1997; 97US-0049556P.  
 PR 13-JUN-1997; 97US-0049557P.  
 PR 13-JUN-1997; 97US-0049558P.  
 PR 13-JUN-1997; 97US-0049559P.  
 PR 13-JUN-1997; 97US-0049560P.  
 PR 13-JUN-1997; 97US-0049561P.  
 PR 13-JUN-1997; 97US-0049562P.  
 PR 13-JUN-1997; 97US-0049563P.  
 PR 13-JUN-1997; 97US-0049564P.  
 PR 13-JUN-1997; 97US-0049565P.  
 PR 13-JUN-1997; 97US-0049566P.  
 PR 13-JUN-1997; 97US-0049567P.  
 PR 13-JUN-1997; 97US-0049568P.  
 PR 13-JUN-1997; 97US-0049569P.  
 PR 13-JUN-1997; 97US-0049570P.  
 PR 13-JUN-1997; 97US-0049571P.  
 PR 13-JUN-1997; 97US-0049572P.  
 PR 13-JUN-1997; 97US-0049573P.  
 PR 13-JUN-1997; 97US-0049574P.  
 PR 13-JUN-1997; 97US-0049575P.  
 PR 13-JUN-1997; 97US-0049576P.  
 PR 13-JUN-1997; 97US-0049577P.  
 PR 13-JUN-1997; 97US-0049578P.  
 PR 13-JUN-1997; 97US-0049579P.  
 PR 13-JUN-1997; 97US-0049580P.  
 PR 13-JUN-1997; 97US-0049581P.  
 PR 13-JUN-1997; 97US-0049582P.  
 PR 13-JUN-1997; 97US-0049583P.  
 PR 13-JUN-1997; 97US-0049584P.  
 PR 13-JUN-1997; 97US-0049585P.  
 PR 13-JUN-1997; 97US-0049586P.  
 PR 13-JUN-1997; 97US-0049587P.  
 PR 13-JUN-1997; 97US-0049588P.  
 PR 13-JUN-1997; 97US-0049589P.  
 PR 13-JUN-1997; 97US-0049590P.  
 PR 13-JUN-1997; 97US-0049591P.  
 PR 13-JUN-1997; 97US-0049592P.  
 PR 13-JUN-1997; 97US-0049593P.  
 PR 13-JUN-1997; 97US-0049594P.  
 PR 13-JUN-1997; 97US-0049595P.  
 PR 13-JUN-1997; 97US-0049596P.  
 PR 13-JUN-1997; 97US-0049597P.  
 PR 13-JUN-1997; 97US-0049598P.  
 PR 13-JUN-1997; 97US-0049599P.  
 PR 13-JUN-1997; 97US-0049600P.  
 PR 13-JUN-1997; 97US-0049601P.  
 PR 13-JUN-1997; 97US-0049602P.  
 PR 13-JUN-1997; 97US-0049603P.  
 PR 13-JUN-1997; 97US-0049604P.  
 PR 13-JUN-1997; 97US-0049605P.  
 PR 13-JUN-1997; 97US-0049606P.  
 PR 13-JUN-1997; 97US-0049607P.  
 PR 13-JUN-1997; 97US-0049608P.  
 PR 13-JUN-1997; 97US-0049609P.  
 PR 13-JUN-1997; 97US-0049610P.  
 PR 13-JUN-1997; 97US-0049611P.  
 PR 13-JUN-1997; 97US-0049612P.  
 PR 13-JUN-1997; 97US-0049613P.  
 PR 13-JUN-1997; 97US-0049614P.  
 PR 13-JUN-1997; 97US-0049615P.  
 PR 13-JUN-1997; 97US-0049616P.  
 PR 13-JUN-1997; 97US-0049617P.  
 PR 13-JUN-1997; 97US-0049618P.  
 PR 13-JUN-1997; 97US-0049619P.  
 PR 13-JUN-1997; 97US-0049620P.  
 PR 13-JUN-1997; 97US-0049621P.  
 PR 13-JUN-1997; 97US-0049622P.  
 PR 13-JUN-1997; 97US-0049623P.  
 PR 13-JUN-1997; 97US-0049624P.  
 PR 13-JUN-1997; 97US-0049625P.  
 PR 13-JUN-1997; 97US-0049626P.  
 PR 13-JUN-1997; 97US-0049627P.  
 PR 13-JUN-1997; 97US-0049628P.  
 PR 13-JUN-1997; 97US-0049629P.  
 PR 13-JUN-1997; 97US-0049630P.  
 PR 13-JUN-1997; 97US-0049631P.  
 PR 13-JUN-1997; 97US-0049632P.  
 PR 13-JUN-1997; 97US-0049633P.  
 PR 13-JUN-1997; 97US-0049634P.  
 PR 13-JUN-1997; 97US-0049635P.  
 PR 13-JUN-1997; 97US-0049636P.  
 PR 13-JUN-1997; 97US-0049637P.  
 PR 13-JUN-1997; 97US-0049638P.  
 PR 13-JUN-1997; 97US-0049639P.  
 PR 13-JUN-1997; 97US-0049640P.  
 PR 13-JUN-1997; 97US-0049641P.  
 PR 13-JUN-1997; 97US-0049642P.  
 PR 13-JUN-1997; 97US-0049643P.  
 PR 13-JUN-1997; 97US-0049644P.  
 PR 13-JUN-1997; 97US-0049645P.  
 PR 13-JUN-1997; 97US-0049646P.  
 PR 13-JUN-1997; 97US-0049647P.  
 PR 13-JUN-1997; 97US-0049648P.  
 PR 13-JUN-1997; 97US-0049649P.  
 PR 13-JUN-1997; 97US-0049650P.  
 PR 13-JUN-1997; 97US-0049651P.  
 PR 13-JUN-1997; 97US-0049652P.  
 PR 13-JUN-1997; 97US-0049653P.  
 PR 13-JUN-1997; 97US-0049654P.  
 PR 13-JUN-1997; 97US-0049655P.  
 PR 13-JUN-1997; 97US-0049656P.  
 PR 13-JUN-1997; 97US-0049657P.  
 PR 13-JUN-1997; 97US-0049658P.  
 PR 13-JUN-1997; 97US-0049659P.  
 PR 13-JUN-1997; 97US-0049660P.  
 PR 13-JUN-1997; 97US-0049661P.  
 PR 13-JUN-1997; 97US-0049662P.  
 PR 13-JUN-1997; 97US-0049663P.  
 PR 13-JUN-1997; 97US-0049664P.  
 PR 13-JUN-1997; 97US-0049665P.  
 PR 13-JUN-1997; 97US-0049666P.  
 PR 13-JUN-1997; 97US-0049667P.  
 PR 13-JUN-1997; 97US-0049668P.  
 PR 13-JUN-1997; 97US-0049669P.  
 PR 13-JUN-1997; 97US-0049670P.  
 PR 13-JUN-1997; 97US-0049671P.  
 PR 13-JUN-1997; 97US-0049672P.  
 PR 13-JUN-1997; 97US-0049673P.  
 PR 13-JUN-1997; 97US-0049674P.  
 PR 13-JUN-1997; 97US-0049675P.  
 PR 13-JUN-1997; 97US-0049676P.  
 PR 13-JUN-1997; 97US-0049677P.  
 PR 13-JUN-1997; 97US-0049678P.  
 PR 13-JUN-1997; 97US-0049679P.  
 PR 13-JUN-1997; 97US-0049680P.  
 PR 13-JUN-1997; 97US-0049681P.  
 PR 13-JUN-1997; 97US-0049682P.  
 PR 13-JUN-1997; 97US-0049683P.  
 PR 13-JUN-1997; 97US-0049684P.  
 PR 13-JUN-1997; 97US-0049685P.  
 PR 13-JUN-1997; 97US-0049686P.  
 PR 13-JUN-1997; 97US-0049687P.  
 PR 13-JUN-1997; 97US-0049688P.  
 PR 13-JUN-1997; 97US-0049689P.  
 PR 13-JUN-1997; 97US-0049690P.  
 PR 13-JUN-1997; 97US-0049691P.  
 PR 13-JUN-1997; 97US-0049692P.  
 PR 13-JUN-1997; 97US-0049693P.  
 PR 13-JUN-1997; 97US-0049694P.  
 PR 13-JUN-1997; 97US-0049695P.  
 PR 13-JUN-1997; 97US-0049696P.  
 PR 13-JUN-1997; 97US-0049697P.  
 PR 13-JUN-1997; 97US-0049698P.  
 PR 13-JUN-1997; 97US-0049699P.  
 PR 13-JUN-1997; 97US-0049700P.  
 PR 13-JUN-1997; 97US-0049701P.  
 PR 13-JUN-1997; 97US-0049702P.  
 PR 13-JUN-1997; 97US-0049703P.  
 PR 13-JUN-1997; 97US-0049704P.  
 PR 13-JUN-1997; 97US-0049705P.  
 PR 13-JUN-1997; 97US-0049706P.  
 PR 13-JUN-1997; 97US-0049707P.  
 PR 13-JUN-1997; 97US-0049708P.  
 PR 13-JUN-1997; 97US-0049709P.  
 PR 13-JUN-1997; 97US-0049710P.  
 PR 13-JUN-1997; 97US-0049711P.  
 PR 13-JUN-1997; 97US-0049712P.  
 PR 13-JUN-1997; 97US-0049713P.  
 PR 13-JUN-1997; 97US-0049714P.  
 PR 13-JUN-1997; 97US-0049715P.  
 PR 13-JUN-1997; 97US-0049716P.  
 PR 13-JUN-1997; 97US-0049717P.  
 PR 13-JUN-1997; 97US-0049718P.  
 PR 13-JUN-1997; 97US-0049719P.  
 PR 13-JUN-1997; 97US-0049720P.  
 PR 13-JUN-1997; 97US-0049721P.  
 PR 13-JUN-1997; 97US-0049722P.  
 PR 13-JUN-1997; 97US-0049723P.  
 PR 13-JUN-1997; 97US-0049724P.  
 PR 13-JUN-1997; 97US-0049725P.  
 PR 13-JUN-1997; 97US-0049726P.  
 PR 13-JUN-1997; 97US-0049727P.  
 PR 13-JUN-1997; 97US-0049728P.  
 PR 13-JUN-1997; 97US-0049729P.  
 PR 13-JUN-1997; 97US-0049730P.  
 PR 13-JUN-1997; 97US-0049731P.  
 PR 13-JUN-1997; 97US-0049732P.  
 PR 13-JUN-1997; 97US-0049733P.  
 PR 13-JUN-1997; 97US-0049734P.  
 PR 13-JUN-1997; 97US-0049735P.  
 PR 13-JUN-1997; 97US-0049736P.  
 PR 13-JUN-1997; 97US-0049737P.  
 PR 13-JUN-1997; 97US-0049738P.  
 PR 13-JUN-1997; 97US-0049739P.  
 PR 13-JUN-1997; 97US-0049740P.  
 PR 13-JUN-1997; 97US-0049741P.  
 PR 13-JUN-1997; 97US-0049742P.  
 PR 13-JUN-1997; 97US-0049743P.  
 PR 13-JUN-1997; 97US-0049744P.  
 PR 13-JUN-1997; 97US-0049745P.  
 PR 13-JUN-1997; 97US-0049746P.  
 PR 13-JUN-1997; 97US-0049747P.  
 PR 13-JUN-1997; 97US-0049748P.  
 PR 13-JUN-1997; 97US-0049749P.  
 PR 13-JUN-1997; 97US-0049750P.  
 PR 13-JUN-1997; 97US-0049751P.  
 PR 13-JUN-1997; 97US-0049752P.  
 PR 13-JUN-1997; 97US-0049753P.  
 PR 13-JUN-1997; 97US-0049754P.  
 PR 13-JUN-1997; 97US-0049755P.  
 PR 13-JUN-1997; 97US-0049756P.  
 PR 13-JUN-1997; 97US-0049757P.  
 PR 13-JUN-1997; 97US-0049758P.  
 PR 13-JUN-1997; 97US-0049759P.  
 PR 13-JUN-1997; 97US-0049760P.  
 PR 13-JUN-1997; 97US-0049761P.  
 PR 13-JUN-1997; 97US-0049762P.  
 PR 13-JUN-1997; 97US-0049763P.  
 PR 13-JUN-1997; 97US-0049764P.  
 PR 13-JUN-1997; 97US-0049765P.  
 PR 13-JUN-1997; 97US-0049766P.  
 PR 13-JUN-1997; 97US-0049767P.  
 PR 13-JUN-1997; 97US-0049768P.  
 PR 13-JUN-1997; 97US-0049769P.  
 PR 13-JUN-1997; 97US-0049770P.  
 PR 13-JUN-1997; 97US-0049771P.  
 PR 13-JUN-1997; 97US-0049772P.  
 PR 13-JUN-1997; 97US-0049773P.  
 PR 13-JUN-1997; 97US-0049774P.  
 PR 13-JUN-1997; 97US-0049775P.  
 PR 13-JUN-1997; 97US-0049776P.  
 PR 13-JUN-1997; 97US-0049777P.  
 PR 13-JUN-1997; 97US-0049778P.  
 PR 13-JUN-1997; 97US-0049779P.  
 PR 13-JUN-1997; 97US-0049780P.  
 PR 13-JUN-1997; 97US-0049781P.  
 PR 13-JUN-1997; 97US-0049782P.  
 PR 13-JUN-1997; 97US-0049783P.  
 PR 13-JUN-1997; 97US-0049784P.  
 PR 13-JUN-1997; 97US-0049785P.  
 PR 13-JUN-1997; 97US-0049786P.  
 PR 13-JUN-1997; 97US-0049787P.  
 PR 13-JUN-1997; 97US-0049788P.  
 PR 13-JUN-1997; 97US-0049789P.  
 PR 13-JUN-1997; 97US-0049790P.  
 PR 13-JUN-1997; 97US-0049791P.  
 PR 13-JUN-1997; 97US-0049792P.  
 PR 13-JUN-1997; 97US-0049793P.  
 PR 13-JUN-1997; 97US-0049794P.  
 PR 13-JUN-1997; 97US-0049795P.  
 PR 13-JUN-1997; 97US-0049796P.  
 PR 13-JUN-1997; 97US-0049797P.  
 PR 13-JUN-1997; 97US-0049798P.  
 PR 13-JUN-1997; 97US-0049799P.  
 PR 13-JUN-1997; 97US-0049800P.  
 PR 13-JUN-1997; 97US-0049801P.  
 PR 13-JUN-1997; 97US-0049802P.  
 PR 13-JUN-1997; 97US-0049803P.  
 PR 13-JUN-1997; 97US-0049804P.  
 PR 13-JUN-1997; 97US-0049805P.  
 PR 13-JUN-1997; 97US-0049806P.  
 PR 13-JUN-1997; 97US-0049807P.  
 PR 13-JUN-1997; 97US-0049808P.  
 PR 13-JUN-1997; 97US-0049809P.  
 PR 13-JUN-1997; 97US-0049810P.  
 PR 13-JUN-1997; 97US-0049811P.  
 PR 13-JUN-1997; 97US-0049812P.  
 PR 13-JUN-1997; 97US-0049813P.  
 PR 13-JUN-1997; 97US-0049814P.  
 PR 13-JUN-1997; 97US-0049815P.  
 PR 13-JUN-1997; 97US-0049816P.  
 PR 13-JUN-1997; 97US-0049817P.  
 PR 13-JUN-1997; 97US-0049818P.  
 PR 13-JUN-1997; 97US-0049819P.  
 PR 13-JUN-1997; 97US-0049820P.  
 PR 13-JUN-1997; 97US-0049821P.  
 PR 13-JUN-1997; 97US-0049822P.  
 PR 13-JUN-1997; 97US-0049823P.  
 PR 13-JUN-1997; 97US-0049824P.  
 PR 13-JUN-1997; 97US-0049825P.  
 PR 13-JUN-1997; 97US-0049826P.  
 PR 13-JUN-1997; 97US-0049827P.  
 PR 13-JUN-1997; 97US-0049828P.  
 PR 13-JUN-1997; 97US-0049829P.  
 PR 13-JUN-1997; 97US-0049830P.  
 PR 13-JUN-1997; 97US-0049831P.  
 PR 13-JUN-1997; 97US-0049832P.  
 PR 13-JUN-1997; 97US-0049833P.  
 PR 13-JUN-1997; 97US-0049834P.  
 PR 13-JUN-1997; 97US-0049835P.  
 PR 13-JUN-1997; 97US-0049836P.  
 PR 13-JUN-1997; 97US-0049837P.  
 PR 13-JUN-1997; 97US-0049838P.  
 PR 13-JUN-1997; 97US-0049839P.  
 PR 13-JUN-1997; 97US-0049840P.  
 PR 13-JUN-1997; 97US-0049841P.  
 PR 13-JUN-1997; 97US-0049842P.  
 PR 13-JUN-1997; 97US-0049843P.  
 PR 13-JUN-1997; 97US-0049844P.  
 PR 13-JUN-1997; 97US-0049845P.  
 PR 13-JUN-1997; 97US-0049846P.  
 PR 13-JUN-1997; 97US-0049847P.  
 PR 13-JUN-1997; 97US-0049848P.  
 PR 13-JUN-1997; 97US-0049849P.  
 PR 13-JUN-1997; 97US-0049850P.  
 PR 13-JUN-1997; 97US-0049851P.  
 PR 13-JUN-1997; 97US-0049852P.  
 PR 13-JUN-1997; 97US-0049853P.  
 PR 13-JUN-1997; 97US-0049854P.  
 PR 13-JUN-1997; 97US-0049855P.  
 PR 13-JUN-1997; 97US-0049856P.  
 PR 13-JUN-1997; 97US-0049857P.  
 PR 13-JUN-1997; 97US-0049858P.  
 PR 13-JUN-1997; 97US-0049859P.  
 PR 13-JUN-1997; 97US-0049860P.  
 PR 13-JUN-1997; 97US-0049861P.  
 PR 13-JUN-1997; 97US-0049862P.  
 PR 13-JUN-1997; 97US-0049863P.  
 PR 13-JUN-1997; 97US-0049864P.  
 PR 13-JUN-1997; 97US-0049865P.  
 PR 13-JUN-1997; 97US-0049866P.  
 PR 13-JUN-1997; 97US-0049867P.  
 PR 13-JUN-1997; 97US-0049868P.  
 PR 13-JUN-1997; 97US-0049869P.  
 PR 13-JUN-1997; 97US-0049870P.  
 PR 13-JUN-1997; 97US-0049871P.  
 PR 13-JUN-1997; 97US-0049872P.  
 PR 13-JUN-1997; 97US-0049873P.  
 PR 13-JUN-1997; 97US-0049874P.  
 PR 13-JUN-1997; 97US-0049875P.  
 PR 13-JUN-1997; 97US-0049876P.  
 PR 13-JUN-1997; 97US-0049877P.  
 PR 13-JUN-1997; 97US-0049878P.  
 PR 13-JUN-1997; 97US-0049879P.  
 PR 13-JUN-1997; 97US-0049880P.  
 PR 13-JUN-1997; 97US-0049881P.  
 PR 13-JUN-1997; 97US-0049882P.  
 PR 13-JUN-1997; 97US-0049883P.  
 PR 13-JUN-1997; 97US-0049884P.  
 PR 13-JUN-1997; 97US-0049885P.  
 PR 13-JUN-1997; 97US-0049886P.  
 PR 13-JUN-1997; 97US-0049887P.  
 PR 13-JUN-1997; 97US-0049888P.  
 PR 13-JUN-1997; 97US-0049889P.  
 PR 13-JUN-1997; 97US-0049890P.  
 PR 13-JUN-1997; 97US-0049891P.  
 PR 13-JUN-1997; 97US-0049892P.  
 PR 13-JUN-1997; 97US-0049893P.  
 PR 13-JUN-1997; 97US-0049894P.  
 PR 13-JUN-1997; 97US-0049895P.  
 PR 13-JUN-1997; 97US-0049896P.  
 PR 13-JUN-1997; 97US-0049897P.  
 PR 13-JUN-1997; 97US-0049898P.  
 PR 13-JUN-1997; 97US-0049899P.  
 PR 13-JUN-1997; 97US-0049900P.  
 PR 13-JUN-1997; 97US-0049901P.  
 PR 13-JUN-1997; 97US-0049902P.  
 PR 13-JUN-1997; 97US-0049903P.  
 PR 13-JUN-1997; 97US-0049904P.  
 PR 13-JUN-1997; 97US-0049905P.  
 PR 13-JUN-1997; 97US-0049906P.  
 PR 13-JUN-1997; 97US-0049907P.  
 PR 13-JUN-1997; 97US-0049908P.  
 PR 13-JUN-1997; 97US-0049909P.  
 PR 13-JUN-1997; 97US-0049910P.  
 PR 13-JUN-1997; 97US-0049911P.  
 PR 13-JUN-1997; 97US-0049912P.  
 PR 1

Db 10 NKESEDKVFLVVKTLQAQLQVTNLLEYLIDVEIARSDCRKPLSTN 55

RESULT 14

ID AD457563 standard; protein; 113 AA.

AC ADA57563;

DT . 20-NOV-2003 (first entry)

DE Human secreted protein #514..

KW immunosuppressive; antiinflammatory; antisthmatic; antiallergic;  
 KW cytosolic; cerebroprotective; neuroprotective; nootropic;  
 KW cardiovascular; antarteriosclerotic; gene therapy;  
 KW human secreted protein; immune disorder; inflammation;  
 KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;  
 KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;  
 KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;  
 KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;  
 KW triple helix formation; antisense gene therapy; forensic biology.

OS Homo sapiens.

PN WO2002102994-A2.

PD 27-DEC-2002.

19-MAR-2002; 2002WO-US0008278.

21-MAR-2001: 2001US-0277340P.

PR 13-NOV-2001: 2001US-0331287P.  
PR 19-JUL-2001: 2001US-0306171P.

XX  
PA (HITMA-) HITMAN GENOME SCT INC

XX	Bogren CA	Burhan SM.
BT		

XX 0000 1234567890

DR N-PSDB; ADA56670.

PT New human secreted p

PT neurodegenerative disorders.  
PT conditions, respiratory disorders, cancers, CNS disorders, or  
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory

PS Claim 13; SEQ ID NO 1756; 1754pp; English.

The invention relates to 592 new human secreted polypeptides useful for diagnosing, treating or preventing e.g. immune disorders, inflammatory conditions, respiratory disorders, cancers, CNS disorders, or neurodegenerative disorders, or polypeptides comprising an amino acid sequence at least 95% identical to the new sequences. The polypeptides, antibodies or antibody fragments that bind to the polypeptides, nucleic acids encoding the polypeptides, agonists or antagonists that binds to the polypeptide, are useful in preparing diagnostic or pharmaceutical compositions for diagnosing, treating or preventing an e.g. immune disorders, inflammatory conditions (e.g. inflammatory bowel disease, nephritis or Crohn's disease), respiratory disorders (e.g. asthma and allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative disorders (e.g. Parkinson's disease or Alzheimer's disease), and cardiovascular disorders (e.g. atherosclerosis or myocarditis). The polynucleotides are useful for chromosome identification, chromosome mapping, for controlling gene expression through triple helix formation or antisense DNA or RNA, in gene therapy, for identifying individuals from minute biological samples, in forensic biology, and as hybridization probes. The polypeptides are useful for as molecular weight markers on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, to raise antibodies, for testing biological activities, and for treating or preventing neural disorders, immune system disorders,

muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, proliferative and/or cancerous diseases. This sequence corresponds to one of the polypeptide of the invention. Note: The sequence data for this patent did form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

**SQ** Sequence 113 AA;

Query Match	46.5%	Score 118	DB 6	Length 113
Best Local Similarity	50.0%	Pred. No. 1.1e-08		
Matches 23	Conservative 11	Mismatches 12	Indels 0	Gaps 0

```
QY      1 NKESDKXHFRIFRVLKVGROVTDHLEXYHLNVENMQWTTQCKPETTN 46
        |||:||| : : | : ||| : ||| : : ||| : : ||| : : |||
Db     58 NKESEDKVFLLVKTLQAOLQVNTNLEYLIDIVEIARSDCRKPLSTN 103
```

Db 58 NKESEDKYVFLVVKTLQAQLQVTNLLLEYLLDVEIARSDCRKPLSTN 103

## RESULT 15

ID	ADA41457	standard; protein; 113 AA
ADA41457	standard; protein; 113 AA	

ADA41457; AC

DT 20-NOV-2003 (first entry)

Human secreted protein.

Human, secreted protein, cancer; hyperproliferative disorder;  
rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;  
anaemia; allergic reaction; asthma; cardiovascular disorder;  
wound healing; cytostatic; immunosuppressive; noctropic; neuroprotective;  
antiviral; antiallergic; hepatotrophic; antidiabetic; antiinflammatory;  
vulnery; cardiant; gene therapy

Homo sapiens.

AA WO2002102993-A2  
PN

27-DEC-2002

XX 19-MAR-2002 . 2002W0-115008123  
PE

XX  
 21 MAY 2001 0001116 0237340D

PR 19-JUL-2001; 2001US-0306171P.

XX

2 XX

XX

XX

PT New human secreted proteins and nucleic acid molecules, useful for preparing a diagnostic or pharmaceutical composition for diagnosing, preventing or treating cancer or other hyperproliferative disorder, PT asthma, allergies or AIDS.

PS Claim 1; SEQ ID NO 1840; 3205pp; English.

CC The invention relates to novel genes AD339623-ADA0565 and proteins  
CC ADA0566-ADA1501 for human secreted proteins, useful for preventing,  
CC treating or ameliorating medical conditions e.g. by protein or gene  
CC therapy. The polypeptides, nucleic acid molecules, antibodies or the  
CC fragments, and agonists or antagonists that bind to the polypeptide are  
CC useful for preparing a diagnostic or pharmaceutical composition for  
CC diagnosing or treating cancer or other hyperproliferative disorder. The  
CC polypeptides and nucleic acid molecules are also useful for detecting,  
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer  
CC or other hyperproliferative disorders including neoplasms, autoimmune  
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus  
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic  
CC anaemia), haemopoietic or haematological disorders (e.g. anaemia,  
CC





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 11.7738 Seconds  
(without alignments)  
201.701 Million cell updates/sec

Title: US-09-941-314-10

Perfect score: 254  
Sequence: 1 NKESSDKYHFRIRVLKQVR.....EYHLNVMQWTCCKPETTN 46

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCUS COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	46.5	142	3	US-09-431-480-4
2	118	46.5	142	3	US-09-617-302-4
3	113	44.5	141	3	US-09-431-480-2
4	113	44.5	141	3	US-09-617-302-2
5	111	43.7	142	3	US-09-431-480-3
6	111	43.7	142	3	US-09-617-302-3
7	103	40.6	127	4	US-08-849-303-19
8	99	39.0	140	4	US-09-886-319A-46
9	99	39.0	140	4	US-09-886-319A-48
10	98	38.6	112	4	US-08-849-303-16
11	98	38.6	118	4	US-09-775-932-24
12	95	37.4	140	3	US-09-431-480-5
13	95	37.4	140	3	US-09-617-302-5
14	95	37.4	140	4	US-08-849-303-18
15	94	37.0	116	4	US-09-775-932-16
16	94	37.0	120	4	US-09-775-932-2
17	94	37.0	120	4	US-09-775-932-2
18	94	37.0	120	6	US-08-791-522-4
19	94	37.0	139	3	US-09-314-777-4
20	94	37.0	139	4	US-08-849-303-15
21	94	37.0	145	2	US-08-832-535-11
22	94	37.0	146	2	US-08-791-522-3
23	94	37.0	146	3	US-08-744-138-3
24	94	37.0	146	3	US-09-019-485-4
25	94	37.0	146	3	US-09-314-777-3
26	94	37.0	146	3	US-09-431-480-6
27	94	37.0	146	3	US-09-617-302-6

28	94	37.0	146	4	US-09-241-376-3	Sequence 3, Appl1
29	94	37.0	146	4	US-09-528-436B-3	Sequence 3, Appl1
30	94	37.0	146	4	US-09-886-319A-47	Sequence 47, Appl1
31	94	37.0	146	4	US-09-940-497-3	Sequence 3, Appl1
32	94	37.0	146	4	US-09-976-594-37	Sequence 3, Appl1
33	94	37.0	146	4	US-08-849-303-17	Sequence 17, Appl1
34	94	37.0	146	4	PCT-US95-07135-9	Sequence 9, Appl1
35	94	37.0	146	6	US-09-775-932-8	Patent No. 5432264
36	84	33.1	121	4	US-09-775-932-8	Sequence 8, Appl1
37	84	33.1	111	3	US-08-744-138-6	Sequence 6, Appl1
38	84	33.1	141	4	US-09-241-376-6	Sequence 6, Appl1
39	84	33.1	141	4	US-09-940-497-6	Sequence 24, Appl1
40	84	33.1	141	4	US-08-849-303-24	Sequence 25, Appl1
41	81	31.9	162	4	US-08-849-303-25	Sequence 25, Appl1
42	80	31.5	111	4	US-08-849-303-26	Sequence 26, Appl1
43	78	30.7	121	4	US-09-775-932-14	Sequence 14, Appl1
44	78	30.7	128	4	US-09-775-932-12	Sequence 12, Appl1
45	78	30.7	149	2	US-08-461-030C-2	Sequence 2, Appl1

## ALIGNMENTS

```

RESULT 1
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Hollaway, James L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; EARLIER FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; NUMBER OF SEQ. ID NOS: 22
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-4

Query Match      46.5%; Score 118; DB 3; Length 142;
Best Local Similarity 50.0%; Pred. No. 4.5e-10;
Matches 23; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Cy      1 NKESSDKYHFRIRVLKQVRQVTDHLEYHLNVMQWTCCKPETTN 46
Db      57 NKESSDKYHFRIRVLKQVRQVTDHLEYHLNVMQWTCCKPETTN 102

RESULT 2
US-09-617-302-4
; Sequence 4, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Feldhaus, Andrew L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; EARLIER FILING DATE: 2000-07-17
; EARLIER APPLICATION NUMBER: 09/431,480
; EARLIER FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22

```

SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-617-302-4

Query Match 46.5%; Score 118; DB 3; Length 142;  
Best Local Similarity 50.0%; Pred. No. 4.5e-10;  
Matches 23; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIFRVLKVGQVTDHLEHNLNEMQWTTCKP 46  
DB 57 NKESDKYVFLVVDILKSDQETDLSLEYLEVINARTWCKK 102

RESULT 3  
US-09-431-480-2  
Sequence 2, Application US/09431480  
Patent No. 6235708  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
APPLICANT: Feldhaus, Andrew  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72  
CURRENT APPLICATION NUMBER: US/09/431,480  
CURRENT FILING DATE: 1999-11-01  
EARLIER APPLICATION NUMBER: 60/109,217  
EARLIER FILING DATE: 1998-11-20  
EARLIER APPLICATION NUMBER: 60/156,382  
EARLIER FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 141  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-431-480-2

Query Match 44.5%; Score 113; DB 3; Length 141;  
Best Local Similarity 51.2%; Pred. No. 2.5e-09;  
Matches 21; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIFRVLKVGQVTDHLEHNLNEMQWTTCK 41  
DB 56 NKASNDLYNFRVVDILKSDQETDLSLEYLEVINARTWCKK 96

RESULT 4  
US-09-617-302-2  
Sequence 2, Application US/09617302  
Patent No. 6245529  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
APPLICANT: Feldhaus, Andrew  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72 C1  
CURRENT APPLICATION NUMBER: US/09/617,302  
CURRENT FILING DATE: 2000-07-17  
PRIOR APPLICATION NUMBER: 09/431,480  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: 60/109,217  
PRIOR FILING DATE: 1998-11-20  
PRIOR APPLICATION NUMBER: 60/156,382  
PRIOR FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 141  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-617-302-2

Query Match 44.5%; Score 113; DB 3; Length 141;  
Best Local Similarity 51.2%; Pred. No. 2.5e-09;  
Matches 21; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIFRVLKVGQVTDHLEHNLNEMQWTTCK 41  
DB 56 NKASNDLYNFRVVDILKSDQETDLSLEYLEVINARTWCKK 96

RESULT 5  
US-09-431-480-3  
Sequence 3, Application US/09431480  
Patent No. 6235708  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
APPLICANT: Feldhaus, Andrew  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72  
CURRENT APPLICATION NUMBER: US/09/431,480  
CURRENT FILING DATE: 1999-11-01  
EARLIER APPLICATION NUMBER: 60/109,217  
EARLIER FILING DATE: 1998-11-20  
EARLIER APPLICATION NUMBER: 60/156,382  
EARLIER FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-431-480-3

Query Match 43.7%; Score 111; DB 3; Length 142;  
Best Local Similarity 42.9%; Pred. No. 5.1e-09;  
Matches 18; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIFRVLKVGQVTDHLEHNLNEMQWTTCKP 42  
DB 57 NKESDKYVFLVVDILKHLAKLQITDREYQIDVOISRSNCKP 98

RESULT 6  
US-09-617-302-3  
Sequence 3, Application US/09617302  
Patent No. 6245529  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
APPLICANT: Feldhaus, Andrew  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72 C1  
CURRENT APPLICATION NUMBER: US/09/617,302  
CURRENT FILING DATE: 2000-07-17  
PRIOR APPLICATION NUMBER: 09/431,480  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: 60/109,217  
PRIOR FILING DATE: 1998-11-20  
PRIOR APPLICATION NUMBER: 60/156,382  
PRIOR FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-617-302-3

Query Match 43.7%; Score 111; DB 3; Length 142;  
Best Local Similarity 42.9%; Pred. No. 5.1e-09;  
Matches 18; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIFRVLKVGQVTDHLEHNLNEMQWTTCKP 42  
DB 57 NKESDKYVFLVVDILKHLAKLQITDREYQIDVOISRSNCKP 98

## RESULT 7

US-08-849-303-19  
Sequence 19, Application US/08849303  
Patent No. 6680424  
GENERAL INFORMATION:  
APPLICANT: Atkinson, Howard J.  
APPLICANT: McPherson, Michael J.  
APPLICANT: Urwin, Peter E.  
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/849,303  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION/DOCKET NUMBER: 26,742  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 127 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-849-303-19

Query Match 40.6%; Score 103; DB 4; Length 127;  
Best Local Similarity 43.2%; Pred. No. 7.1e-08;  
Matches 19; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

Qy 1 NKSDDKTHFRIRVLYKQROVTDHLEHYLNVMQWTTCKPPT 44  
Db 42 NKSNDAYHSRAIQVVRARQOLVAGVNYFLDVEMGRITTKSQGT 85

## RESULT 8

US-09-886-319A-46  
Sequence 46, Application US/09886319A  
Patent No. 6586185  
GENERAL INFORMATION:  
APPLICANT: Wolf, Eckard  
APPLICANT: Werner, Sabine  
APPLICANT: Haile, Jorn-Peter  
APPLICANT: Regenbogen, Johannes  
APPLICANT: Goppelt, Andreas  
TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound  
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically  
TITLE OF INVENTION: Active Substances  
FILE REFERENCE: 50125/014002  
CURRENT APPLICATION NUMBER: US/09/886,319A  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
PRIOR APPLICATION NUMBER: US 60/222,081

PRIOR FILING DATE: 2000-08-01  
PRIOR APPLICATION NUMBER: DE 10030149.5  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 46  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-886-319A-46

Query Match 39.0%; Score 99; DB 4; Length 140;  
Best Local Similarity 43.2%; Pred. No. 3.2e-07;  
Matches 19; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

Qy 1 NKSDDKTHFRIRVLYKQROVTDHLEHYLNVMQWTTCKPPT 44  
Db 55 NKSNDAYHSRAIQVVRARQOLVAGVNYFLDVEMGRITTKSQGT 98

## RESULT 9

US-09-886-319A-48  
Sequence 48, Application US/09886319A  
Patent No. 6586185  
GENERAL INFORMATION:  
APPLICANT: Wolf, Eckard  
APPLICANT: Werner, Sabine  
APPLICANT: Haile, Jorn-Peter  
APPLICANT: Regenbogen, Johannes  
APPLICANT: Goppelt, Andreas  
TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound  
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically  
TITLE OF INVENTION: Active Substances  
FILE REFERENCE: 50125/014002  
CURRENT APPLICATION NUMBER: US/09/886,319A  
CURRENT FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: US 60/222,081  
PRIOR FILING DATE: 2000-08-01  
PRIOR APPLICATION NUMBER: DE 10030149.5  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 48  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-886-319A-48

Query Match 39.0%; Score 99; DB 4; Length 140;  
Best Local Similarity 43.2%; Pred. No. 3.2e-07;  
Matches 19; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

Qy 1 NKSDDKTHFRIRVLYKQROVTDHLEHYLNVMQWTTCKPPT 44  
Db 55 NKSNDAYHSRAIQVVRARQOLVAGVNYFLDVEMGRITTKSQGT 98

## RESULT 10

US-08-849-303-16  
Sequence 16, Application US/08849303  
Patent No. 6680424  
GENERAL INFORMATION:  
APPLICANT: Atkinson, Howard J.  
APPLICANT: McPherson, Michael J.  
APPLICANT: Urwin, Peter E.  
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey

COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/849,303  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1321-1-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
US-08-849-303-16

Query Match 38.6%; Score 98; DB 4; Length 112;  
Best Local Similarity 41.9%; Pred. No. 3.4e-07;  
Matches 18; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRYLVKQROVTDHLEHNLVEMQWTTCKPE 43  
DB 28 NKRSNDAYQSRVVRVVRARQVVGMYFLDVELGRTTCTKSQ 70

RESULT 11  
US-09-775-932-24  
Sequence 24, Application US/09775932  
Patent No. 6534477  
GENERAL INFORMATION:  
APPLICANT: University of British Columbia  
TITLE OF INVENTION: Production and use of Modified Cystatins  
FILE REFERENCE: 58069  
CURRENT APPLICATION NUMBER: US/09/775,932  
CURRENT FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: CA99/00717  
PRIOR FILING DATE: 1999-08-05  
PRIOR APPLICATION NUMBER: 60/095,503  
PRIOR FILING DATE: 1998-08-05  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 24  
LENGTH: 118  
TYPE: PRT  
ORGANISM: Bos taurus  
US-09-775-932-24

Query Match 38.6%; Score 98; DB 4; Length 118;  
Best Local Similarity 41.9%; Pred. No. 3.6e-07;  
Matches 18; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRYLVKQROVTDHLEHNLVEMQWTTCKPE 43  
DB 34 NKRSNDAYQSRVVRVVRARQVVGMYFLDVELGRTTCTKSQ 76

RESULT 12  
US-09-431-480-5  
Sequence 5, Application US/09431480  
Patent No. 6235708

GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
APPLICANT: Feldhaus, Andrew  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72  
CURRENT APPLICATION NUMBER: US/09/431,480  
CURRENT FILING DATE: 1999-11-01  
EARLIER APPLICATION NUMBER: 60/109,217  
EARLIER FILING DATE: 1998-11-20  
EARLIER APPLICATION NUMBER: 60/156,382  
EARLIER FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 5  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-431-480-5

Query Match 37.4%; Score 95; DB 3; Length 140;  
Best Local Similarity 40.9%; Pred. No. 1.3e-06;  
Matches 18; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRYLVKQROVTDHLEHNLVEMQWTTCKPE 44  
DB 55 NKRSNDAYHRSRAIQVVRARQVLVAGVYFPDVEGRTTCTKSQT 98

RESULT 13  
US-09-617-302-5  
Sequence 5, Application US/09617302  
Patent No. 6245529  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72 C1  
CURRENT APPLICATION NUMBER: US/09/617,302  
CURRENT FILING DATE: 2000-07-17  
PRIOR APPLICATION NUMBER: 09/431,480  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: 60/109,217  
PRIOR FILING DATE: 1998-11-20  
PRIOR APPLICATION NUMBER: 60/156,382  
PRIOR FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 5  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-617-302-5

Query Match 37.4%; Score 95; DB 3; Length 140;  
Best Local Similarity 40.9%; Pred. No. 1.3e-06;  
Matches 18; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRYLVKQROVTDHLEHNLVEMQWTTCKPE 44  
DB 55 NKRSNDAYHRSRAIQVVRARQVLVAGVYFPDVEGRTTCTKSQT 98

RESULT 14  
US-08-849-303-18  
Sequence 18, Application US/08849303  
Patent No. 6680424  
GENERAL INFORMATION:  
APPLICANT: Atkinson, Howard J.  
APPLICANT: McPherson, Michael J.  
APPLICANT: Urein, Peter E.  
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:

Search completed: March 18, 2004, 14:25:36  
Job time : 11.7738 secs

ADDRESSEE: Klaubert & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/849,303  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1321-1-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 140 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-849-303-18

Query Match 37.4%; Score 95; DB 4; Length 140;  
Best Local Similarity 40.9%; Pred. No. 1.3e-06;  
Matches 18; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIPRVLVKQROVTDHLEHNLVEMQMTTCCKP 44  
DB 55 NKGSDNAYHSRAIQVRRARQQLVAGVNYFEDVENGRITTCYSQT 98

RESULT 15  
US-09-775-932-16  
; Sequence 16; Application US/09775932  
; Patent No. 6534477  
; GENERAL INFORMATION:  
; APPLICANT: University of British Columbia  
; TITLE OF INVENTION: Production and use of Modified Cytatins  
; FILE REFERENCE: 58069  
; CURRENT APPLICATION NUMBER: US/09/775,932  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: CA99/00717  
; PRIOR FILING DATE: 1999-08-05  
; PRIOR APPLICATION NUMBER: 60/095,503  
; PRIOR FILING DATE: 1998-08-05  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Gallus sp.  
US-09-775-932-16

Query Match 37.0%; Score 94; DB 4; Length 116;  
Best Local Similarity 43.9%; Pred. No. 1.4e-06;  
Matches 18; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIPRVLVKQROVTDHLEHNLVEMQMTTCCK 41  
DB 33 NRASNDKYSRVRVVISAKROLVSGIKYIIQVEIGRTTCRK 73



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 27.8373 Seconds  
(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-10

Perfect score: 254  
Sequence: 1 NKSDDKXHFRIPLVKVOR.....EYHLNVQMWTCKPPTTN 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/2/pubppa/PCRT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/2/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/2/pubppa/PCRTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/2/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
13: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/prodata/2/pubppa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/prodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	100.0	46	9	US-09-941-314-10 Sequence 10, App1
2	254	100.0	49	9	US-09-941-314-13 Sequence 13, App1
3	254	100.0	80	9	US-09-941-314-15 Sequence 15, App1
4	254	100.0	115	9	US-09-941-314-3 Sequence 3, App1
5	254	100.0	117	9	US-09-941-314-4 Sequence 4, App1
6	254	100.0	137	9	US-09-941-314-2 Sequence 2, App1
7	248	97.6	52	9	US-09-941-314-14 Sequence 14, App1
8	154	60.6	59	9	US-09-941-314-16 Sequence 16, App1
9	149	58.7	33	9	US-09-941-314-12 Sequence 12, App1
10	138	54.3	24	9	US-09-941-314-11 Sequence 11, App1
11	125	49.2	27	9	US-09-941-314-9 Sequence 9, App1
12	125	49.2	35	9	US-09-941-314-8 Sequence 8, App1
13	110	43.3	50	9	US-09-864-761-34822 Sequence 34822, A
14	110	43.3	50	9	US-09-864-761-48936 Sequence 48936, A
15	107	42.1	145	9	US-09-740-638-2 Sequence 2, App1

16	107	42.1	145	13	US-10-006-467-2	Sequence 2, App1
17 <td>107<td>42.1<td>145<td>14<td>US-10-235-148-2<td>Sequence 2, App1</td></td></td></td></td></td>	107 <td>42.1<td>145<td>14<td>US-10-235-148-2<td>Sequence 2, App1</td></td></td></td></td>	42.1 <td>145<td>14<td>US-10-235-148-2<td>Sequence 2, App1</td></td></td></td>	145 <td>14<td>US-10-235-148-2<td>Sequence 2, App1</td></td></td>	14 <td>US-10-235-148-2<td>Sequence 2, App1</td></td>	US-10-235-148-2 <td>Sequence 2, App1</td>	Sequence 2, App1
18 <td>103<td>40.6<td>127<td>8<td>US-08-849-303-19<td>Sequence 19, App1</td></td></td></td></td></td>	103 <td>40.6<td>127<td>8<td>US-08-849-303-19<td>Sequence 19, App1</td></td></td></td></td>	40.6 <td>127<td>8<td>US-08-849-303-19<td>Sequence 19, App1</td></td></td></td>	127 <td>8<td>US-08-849-303-19<td>Sequence 19, App1</td></td></td>	8 <td>US-08-849-303-19<td>Sequence 19, App1</td></td>	US-08-849-303-19 <td>Sequence 19, App1</td>	Sequence 19, App1
19 <td>101<td>39.8<td>145<td>14<td>US-10-168-425-14<td>Sequence 14, App1</td></td></td></td></td></td>	101 <td>39.8<td>145<td>14<td>US-10-168-425-14<td>Sequence 14, App1</td></td></td></td></td>	39.8 <td>145<td>14<td>US-10-168-425-14<td>Sequence 14, App1</td></td></td></td>	145 <td>14<td>US-10-168-425-14<td>Sequence 14, App1</td></td></td>	14 <td>US-10-168-425-14<td>Sequence 14, App1</td></td>	US-10-168-425-14 <td>Sequence 14, App1</td>	Sequence 14, App1
20 <td>99<td>39.0<td>140<td>14<td>US-10-376-564-46<td>Sequence 46, App1</td></td></td></td></td></td>	99 <td>39.0<td>140<td>14<td>US-10-376-564-46<td>Sequence 46, App1</td></td></td></td></td>	39.0 <td>140<td>14<td>US-10-376-564-46<td>Sequence 46, App1</td></td></td></td>	140 <td>14<td>US-10-376-564-46<td>Sequence 46, App1</td></td></td>	14 <td>US-10-376-564-46<td>Sequence 46, App1</td></td>	US-10-376-564-46 <td>Sequence 46, App1</td>	Sequence 46, App1
21 <td>99<td>39.0<td>140<td>14<td>US-10-376-564-48<td>Sequence 48, App1</td></td></td></td></td></td>	99 <td>39.0<td>140<td>14<td>US-10-376-564-48<td>Sequence 48, App1</td></td></td></td></td>	39.0 <td>140<td>14<td>US-10-376-564-48<td>Sequence 48, App1</td></td></td></td>	140 <td>14<td>US-10-376-564-48<td>Sequence 48, App1</td></td></td>	14 <td>US-10-376-564-48<td>Sequence 48, App1</td></td>	US-10-376-564-48 <td>Sequence 48, App1</td>	Sequence 48, App1
22 <td>98<td>38.6<td>112<td>8<td>US-08-849-303-16<td>Sequence 16, App1</td></td></td></td></td></td>	98 <td>38.6<td>112<td>8<td>US-08-849-303-16<td>Sequence 16, App1</td></td></td></td></td>	38.6 <td>112<td>8<td>US-08-849-303-16<td>Sequence 16, App1</td></td></td></td>	112 <td>8<td>US-08-849-303-16<td>Sequence 16, App1</td></td></td>	8 <td>US-08-849-303-16<td>Sequence 16, App1</td></td>	US-08-849-303-16 <td>Sequence 16, App1</td>	Sequence 16, App1
23 <td>98<td>38.6<td>118<td>9<td>US-09-775-932-24<td>Sequence 24, App1</td></td></td></td></td></td>	98 <td>38.6<td>118<td>9<td>US-09-775-932-24<td>Sequence 24, App1</td></td></td></td></td>	38.6 <td>118<td>9<td>US-09-775-932-24<td>Sequence 24, App1</td></td></td></td>	118 <td>9<td>US-09-775-932-24<td>Sequence 24, App1</td></td></td>	9 <td>US-09-775-932-24<td>Sequence 24, App1</td></td>	US-09-775-932-24 <td>Sequence 24, App1</td>	Sequence 24, App1
24 <td>98<td>38.6<td>181<td>15<td>US-10-264-043-2608<td>Sequence 2608, App1</td></td></td></td></td></td>	98 <td>38.6<td>181<td>15<td>US-10-264-043-2608<td>Sequence 2608, App1</td></td></td></td></td>	38.6 <td>181<td>15<td>US-10-264-043-2608<td>Sequence 2608, App1</td></td></td></td>	181 <td>15<td>US-10-264-043-2608<td>Sequence 2608, App1</td></td></td>	15 <td>US-10-264-043-2608<td>Sequence 2608, App1</td></td>	US-10-264-043-2608 <td>Sequence 2608, App1</td>	Sequence 2608, App1
25 <td>95<td>37.4<td>110<td>8<td>US-08-849-303-18<td>Sequence 18, App1</td></td></td></td></td></td>	95 <td>37.4<td>110<td>8<td>US-08-849-303-18<td>Sequence 18, App1</td></td></td></td></td>	37.4 <td>110<td>8<td>US-08-849-303-18<td>Sequence 18, App1</td></td></td></td>	110 <td>8<td>US-08-849-303-18<td>Sequence 18, App1</td></td></td>	8 <td>US-08-849-303-18<td>Sequence 18, App1</td></td>	US-08-849-303-18 <td>Sequence 18, App1</td>	Sequence 18, App1
26 <td>94<td>37.0<td>116<td>9<td>US-09-775-932-16<td>Sequence 16, App1</td></td></td></td></td></td>	94 <td>37.0<td>116<td>9<td>US-09-775-932-16<td>Sequence 16, App1</td></td></td></td></td>	37.0 <td>116<td>9<td>US-09-775-932-16<td>Sequence 16, App1</td></td></td></td>	116 <td>9<td>US-09-775-932-16<td>Sequence 16, App1</td></td></td>	9 <td>US-09-775-932-16<td>Sequence 16, App1</td></td>	US-09-775-932-16 <td>Sequence 16, App1</td>	Sequence 16, App1
27 <td>94<td>37.0<td>120<td>9<td>US-09-775-932-2<td>Sequence 2, App1</td></td></td></td></td></td>	94 <td>37.0<td>120<td>9<td>US-09-775-932-2<td>Sequence 2, App1</td></td></td></td></td>	37.0 <td>120<td>9<td>US-09-775-932-2<td>Sequence 2, App1</td></td></td></td>	120 <td>9<td>US-09-775-932-2<td>Sequence 2, App1</td></td></td>	9 <td>US-09-775-932-2<td>Sequence 2, App1</td></td>	US-09-775-932-2 <td>Sequence 2, App1</td>	Sequence 2, App1
28 <td>94<td>37.0<td>139<td>8<td>US-08-849-303-15<td>Sequence 15, App1</td></td></td></td></td></td>	94 <td>37.0<td>139<td>8<td>US-08-849-303-15<td>Sequence 15, App1</td></td></td></td></td>	37.0 <td>139<td>8<td>US-08-849-303-15<td>Sequence 15, App1</td></td></td></td>	139 <td>8<td>US-08-849-303-15<td>Sequence 15, App1</td></td></td>	8 <td>US-08-849-303-15<td>Sequence 15, App1</td></td>	US-08-849-303-15 <td>Sequence 15, App1</td>	Sequence 15, App1
29 <td>94<td>37.0<td>139<td>9<td>US-09-969-834-4<td>Sequence 4, App1</td></td></td></td></td></td>	94 <td>37.0<td>139<td>9<td>US-09-969-834-4<td>Sequence 4, App1</td></td></td></td></td>	37.0 <td>139<td>9<td>US-09-969-834-4<td>Sequence 4, App1</td></td></td></td>	139 <td>9<td>US-09-969-834-4<td>Sequence 4, App1</td></td></td>	9 <td>US-09-969-834-4<td>Sequence 4, App1</td></td>	US-09-969-834-4 <td>Sequence 4, App1</td>	Sequence 4, App1
30 <td>94<td>37.0<td>146<td>8<td>US-08-849-303-17<td>Sequence 17, App1</td></td></td></td></td></td>	94 <td>37.0<td>146<td>8<td>US-08-849-303-17<td>Sequence 17, App1</td></td></td></td></td>	37.0 <td>146<td>8<td>US-08-849-303-17<td>Sequence 17, App1</td></td></td></td>	146 <td>8<td>US-08-849-303-17<td>Sequence 17, App1</td></td></td>	8 <td>US-08-849-303-17<td>Sequence 17, App1</td></td>	US-08-849-303-17 <td>Sequence 17, App1</td>	Sequence 17, App1
31 <td>94<td>37.0<td>146<td>9<td>US-09-940-497-3<td>Sequence 3, App1</td></td></td></td></td></td>	94 <td>37.0<td>146<td>9<td>US-09-940-497-3<td>Sequence 3, App1</td></td></td></td></td>	37.0 <td>146<td>9<td>US-09-940-497-3<td>Sequence 3, App1</td></td></td></td>	146 <td>9<td>US-09-940-497-3<td>Sequence 3, App1</td></td></td>	9 <td>US-09-940-497-3<td>Sequence 3, App1</td></td>	US-09-940-497-3 <td>Sequence 3, App1</td>	Sequence 3, App1
32 <td>94<td>37.0<td>146<td>9<td>US-09-969-834-3<td>Sequence 3, App1</td></td></td></td></td></td>	94 <td>37.0<td>146<td>9<td>US-09-969-834-3<td>Sequence 3, App1</td></td></td></td></td>	37.0 <td>146<td>9<td>US-09-969-834-3<td>Sequence 3, App1</td></td></td></td>	146 <td>9<td>US-09-969-834-3<td>Sequence 3, App1</td></td></td>	9 <td>US-09-969-834-3<td>Sequence 3, App1</td></td>	US-09-969-834-3 <td>Sequence 3, App1</td>	Sequence 3, App1
33 <td>94<td>37.0<td>146<td>14<td>US-10-329-428-3<td>Sequence 3, App1</td></td></td></td></td></td>	94 <td>37.0<td>146<td>14<td>US-10-329-428-3<td>Sequence 3, App1</td></td></td></td></td>	37.0 <td>146<td>14<td>US-10-329-428-3<td>Sequence 3, App1</td></td></td></td>	146 <td>14<td>US-10-329-428-3<td>Sequence 3, App1</td></td></td>	14 <td>US-10-329-428-3<td>Sequence 3, App1</td></td>	US-10-329-428-3 <td>Sequence 3, App1</td>	Sequence 3, App1
34 <td>94<td>37.0<td>146<td>14<td>US-10-376-564-47<td>Sequence 47, App1</td></td></td></td></td></td>	94 <td>37.0<td>146<td>14<td>US-10-376-564-47<td>Sequence 47, App1</td></td></td></td></td>	37.0 <td>146<td>14<td>US-10-376-564-47<td>Sequence 47, App1</td></td></td></td>	146 <td>14<td>US-10-376-564-47<td>Sequence 47, App1</td></td></td>	14 <td>US-10-376-564-47<td>Sequence 47, App1</td></td>	US-10-376-564-47 <td>Sequence 47, App1</td>	Sequence 47, App1
35 <td>93<td>36.6<td>48<td>9<td>US-09-941-314-17<td>Sequence 17, App1</td></td></td></td></td></td>	93 <td>36.6<td>48<td>9<td>US-09-941-314-17<td>Sequence 17, App1</td></td></td></td></td>	36.6 <td>48<td>9<td>US-09-941-314-17<td>Sequence 17, App1</td></td></td></td>	48 <td>9<td>US-09-941-314-17<td>Sequence 17, App1</td></td></td>	9 <td>US-09-941-314-17<td>Sequence 17, App1</td></td>	US-09-941-314-17 <td>Sequence 17, App1</td>	Sequence 17, App1
36 <td>87<td>34.3<td>165<td>9<td>US-09-740-638-5<td>Sequence 5, App1</td></td></td></td></td></td>	87 <td>34.3<td>165<td>9<td>US-09-740-638-5<td>Sequence 5, App1</td></td></td></td></td>	34.3 <td>165<td>9<td>US-09-740-638-5<td>Sequence 5, App1</td></td></td></td>	165 <td>9<td>US-09-740-638-5<td>Sequence 5, App1</td></td></td>	9 <td>US-09-740-638-5<td>Sequence 5, App1</td></td>	US-09-740-638-5 <td>Sequence 5, App1</td>	Sequence 5, App1
37 <td>87<td>34.3<td>165<td>13<td>US-10-006-467-5<td>Sequence 5, App1</td></td></td></td></td></td>	87 <td>34.3<td>165<td>13<td>US-10-006-467-5<td>Sequence 5, App1</td></td></td></td></td>	34.3 <td>165<td>13<td>US-10-006-467-5<td>Sequence 5, App1</td></td></td></td>	165 <td>13<td>US-10-006-467-5<td>Sequence 5, App1</td></td></td>	13 <td>US-10-006-467-5<td>Sequence 5, App1</td></td>	US-10-006-467-5 <td>Sequence 5, App1</td>	Sequence 5, App1
38 <td>87<td>34.3<td>165<td>14<td>US-10-235-148-5<td>Sequence 5, App1</td></td></td></td></td></td>	87 <td>34.3<td>165<td>14<td>US-10-235-148-5<td>Sequence 5, App1</td></td></td></td></td>	34.3 <td>165<td>14<td>US-10-235-148-5<td>Sequence 5, App1</td></td></td></td>	165 <td>14<td>US-10-235-148-5<td>Sequence 5, App1</td></td></td>	14 <td>US-10-235-148-5<td>Sequence 5, App1</td></td>	US-10-235-148-5 <td>Sequence 5, App1</td>	Sequence 5, App1
39 <td>84<td>33.1<td>121</td><td>9<td>US-09-775-932-8<td>Sequence 8, App1</td></td></td></td></td>	84 <td>33.1<td>121</td><td>9<td>US-09-775-932-8<td>Sequence 8, App1</td></td></td></td>	33.1 <td>121</td> <td>9<td>US-09-775-932-8<td>Sequence 8, App1</td></td></td>	121	9 <td>US-09-775-932-8<td>Sequence 8, App1</td></td>	US-09-775-932-8 <td>Sequence 8, App1</td>	Sequence 8, App1
40 <td>84<td>33.1<td>141<td>8<td>US-08-849-303-24<td>Sequence 24, App1</td></td></td></td></td></td>	84 <td>33.1<td>141<td>8<td>US-08-849-303-24<td>Sequence 24, App1</td></td></td></td></td>	33.1 <td>141<td>8<td>US-08-849-303-24<td>Sequence 24, App1</td></td></td></td>	141 <td>8<td>US-08-849-303-24<td>Sequence 24, App1</td></td></td>	8 <td>US-08-849-303-24<td>Sequence 24, App1</td></td>	US-08-849-303-24 <td>Sequence 24, App1</td>	Sequence 24, App1
41 <td>84<td>33.1<td>141<td>9<td>US-09-940-497-6<td>Sequence 6, App1</td></td></td></td></td></td>	84 <td>33.1<td>141<td>9<td>US-09-940-497-6<td>Sequence 6, App1</td></td></td></td></td>	33.1 <td>141<td>9<td>US-09-940-497-6<td>Sequence 6, App1</td></td></td></td>	141 <td>9<td>US-09-940-497-6<td>Sequence 6, App1</td></td></td>	9 <td>US-09-940-497-6<td>Sequence 6, App1</td></td>	US-09-940-497-6 <td>Sequence 6, App1</td>	Sequence 6, App1
42 <td>81<td>31.9<td>162<td>8<td>US-08-849-303-25<td>Sequence 25, App1</td></td></td></td></td></td>	81 <td>31.9<td>162<td>8<td>US-08-849-303-25<td>Sequence 25, App1</td></td></td></td></td>	31.9 <td>162<td>8<td>US-08-849-303-25<td>Sequence 25, App1</td></td></td></td>	162 <td>8<td>US-08-849-303-25<td>Sequence 25, App1</td></td></td>	8 <td>US-08-849-303-25<td>Sequence 25, App1</td></td>	US-08-849-303-25 <td>Sequence 25, App1</td>	Sequence 25, App1
43 <td>80<td>31.5<td>111<td>8<td>US-08-849-303-26<td>Sequence 26, App1</td></td></td></td></td></td>	80 <td>31.5<td>111<td>8<td>US-08-849-303-26<td>Sequence 26, App1</td></td></td></td></td>	31.5 <td>111<td>8<td>US-08-849-303-26<td>Sequence 26, App1</td></td></td></td>	111 <td>8<td>US-08-849-303-26<td>Sequence 26, App1</td></td></td>	8 <td>US-08-849-303-26<td>Sequence 26, App1</td></td>	US-08-849-303-26 <td>Sequence 26, App1</td>	Sequence 26, App1
44 <td>78<td>30.7<td>121</td><td>9<td>US-09-775-932-14<td>Sequence 14, App1</td></td></td></td></td>	78 <td>30.7<td>121</td><td>9<td>US-09-775-932-14<td>Sequence 14, App1</td></td></td></td>	30.7 <td>121</td> <td>9<td>US-09-775-932-14<td>Sequence 14, App1</td></td></td>	121	9 <td>US-09-775-932-14<td>Sequence 14, App1</td></td>	US-09-775-932-14 <td>Sequence 14, App1</td>	Sequence 14, App1
45 <td>78<td>30.7<td>128<td>9<td>US-09-775-932-12<td>Sequence 12, App1</td></td></td></td></td></td>	78 <td>30.7<td>128<td>9<td>US-09-775-932-12<td>Sequence 12, App1</td></td></td></td></td>	30.7 <td>128<td>9<td>US-09-775-932-12<td>Sequence 12, App1</td></td></td></td>	128 <td>9<td>US-09-775-932-12<td>Sequence 12, App1</td></td></td>	9 <td>US-09-775-932-12<td>Sequence 12, App1</td></td>	US-09-775-932-12 <td>Sequence 12, App1</td>	Sequence 12, App1

## ALIGNMENTS

RESULT 1  
US-09-941-314-10  
Sequence 10, Application US/09941314  
Patient No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 46  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-10

Query Match 100.0%; Score 254; DB 9; Length 46;  
Best Local Similarity 100.0%; Pred. No. 2, 2e-27;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKSDDKXHFRIPLVKVORQVTDHLEHYHLNVQMWTCKPPTTN 46  
DB 1 NKSDDKXHFRIPLVKVORQVTDHLEHYHLNVQMWTCKPPTTN 46

RESULT 2  
US-09-941-314-13  
Sequence 13, Application US/09941314  
Patient No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
Inhibit Cancer Procoagulant Protein

FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-13

Query Match 100.0%; Score 254; DB 9; Length 49;  
Best Local Similarity 100.0%; Pred. No. 2.4e-27;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46  
DB 4 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 49

RESULT 3  
US-09-941-314-15  
Sequence 15, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-15

Query Match 100.0%; Score 254; DB 9; Length 80;  
Best Local Similarity 100.0%; Pred. No. 4.2e-27;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46  
DB 3 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 48

RESULT 4  
US-09-941-314-3  
Sequence 3, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 115  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-3

Query Match 100.0%; Score 254; DB 9; Length 115;  
Best Local Similarity 100.0%; Pred. No. 6.4e-27;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46  
DB 33 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 78

RESULT 5  
US-09-941-314-4  
Sequence 4, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 117  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-4

Query Match 100.0%; Score 254; DB 9; Length 117;  
Best Local Similarity 100.0%; Pred. No. 6.5e-27;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46  
DB 35 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 80

RESULT 6  
US-09-941-314-2  
Sequence 2, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 137  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-2

Query Match 100.0%; Score 254; DB 9; Length 137;  
Best Local Similarity 100.0%; Pred. No. 7.7e-27;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46  
DB 55 NKSDDKXHFRIFRVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 100

RESULT 7  
US-09-941-314-14  
Sequence 14, Application US/09941314  
Patent No. US20020142396A1

```
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-14

Query Match      97.6%; Score 248; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 1,7e-26;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 KESDDKYHFRIRPVLYKQROVTDHLEHYHLNVEMQWTTCKRPETN 46
DB      1 KESDDKYHFRIRPVLYKQROVTDHLEHYHLNVEMQWTTCKRPETN 45

RESULT 8
US-09-941-314-16
; Sequence 16, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-16

Query Match      60.6%; Score 154; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 1,4e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 RQVTDHLEHYHLNVEMQWTTCKRPETN 46
DB      1 RQVTDHLEHYHLNVEMQWTTCKRPETN 27

RESULT 9
US-09-941-314-12
; Sequence 12, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 33
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-12

Query Match      58.7%; Score 149; DB 9; Length 33;
Best Local Similarity 100.0%; Pred. No. 3,6e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      21 QVTDHLEHYHLNVEMQWTTCKRPETN 46
DB      1 QVTDHLEHYHLNVEMQWTTCKRPETN 26

RESULT 10
US-09-941-314-11
; Sequence 11, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-11

Query Match      54.3%; Score 138; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 8e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 RQVTDHLEHYHLNVEMQWTTCKRPE 43
DB      1 RQVTDHLEHYHLNVEMQWTTCKRPE 24

RESULT 11
US-09-941-314-9
; Sequence 9, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-9

Query Match      49.2%; Score 125; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 5,5e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NKESDDKYHFRIRPVLYKQROVTD 24
DB      4 NKESDDKYHFRIRPVLYKQROVTD 27
```

```

RESULT 12
US-09-941-314-8
Sequence 8, Application US/09941314
Patent No. US200200142396A1
GENERAL INFORMATION:
APPLICANT: Zymogenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8
TITLE OF INVENTION: Inhibit Cancer Proco
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09-941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatsSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 35
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-8

```

Query Match	49.2%	Score 125;	DB 9;	Length 35;
Best Local Similarity	100.0%	Pred. No. 7.4e-10;		
Matches 24;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1 NKSDDKYHFRIFRLKVGQRTD 24
        |||||
Db      12 NKSDDKYHFRIFRLKVGQRTD 35
```

RESULT 13  
US-09-864-761-34822

Sequence 34822, Application US/03064761  
Patent No. US20020040763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```

1      PRIOR FILING DATE: 2001-01-30
2      PRIOR APPLICATION NUMBER: US 60/234,687
3      PRIOR FILING DATE: 2000-09-21
4      PRIOR APPLICATION NUMBER: US 09/608,408
5      PRIOR FILING DATE: 2000-06-30
6      PRIOR APPLICATION NUMBER: US 09/774,203
7      PRIOR FILING DATE: 2001-01-29
8      NUMBER OF SEQ ID NOS: 49117
9      SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
10     SEQ ID NO 34822
11     TYPE: PR1
12     LENGTH: 50
13     ORGANISM: Homo sapiens
14     FEATURE:
15     OTHER INFORMATION: MAP TO AL109954.10
16     OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
17     OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
18     OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
19     OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
20     OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
21     OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
22     OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
23     OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
24     OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
25     OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
26     OTHER INFORMATION: EST HUMAN HIT: A120857.1, EVALUE 5.00e-23
27     OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUE 1.00e-01
28     US-09-864-761-34822

```

Query Match	43.3%	Score 110;	DB 9;	Length 50;
Best Local Similarity	100.0%;	Pred. No. 1.2e-07;		
Matches 21; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1 NKESSDKYHFRIFRVLKVQRQ 21  
        |||||  
DB     30 NKESSDKYHFRIFRVLKVQRQ 50
```

```

RESULT 14
US-09-864-761-48936
: Sequence 48936, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Mengsheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecmica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668

```

Fri Mar 19 07:21:53 2004

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48936
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL096677.18
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96
; OTHER INFORMATION: EST HUMAN HIT: A120857.1, EVALUE 5.00e-23
; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUE 1.00e-01
US-09-864-761-48936

```

```

Query Match          43.3%; Score 110; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 NKSDDKYHFRIRPVLYKVQRQ 21
Db      30 NKSDDKYHFRIRPVLYKVQRQ 50

```

```

RESULT 15
US-09-740-638-2
; Sequence 2, Application US/09740638
; Patent No. US20020006656A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Zcy95: A Member of the Cystatin
; FILE REFERENCE: 99-104
; CURRENT APPLICATION NUMBER: US/09/740,638
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-740-638-2

```

```

Query Match          42.1%; Score 107; DB 9; Length 145;
Best Local Similarity 39.1%; Pred. No. 1.1e-06;
Matches 18; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

```

```

QY      1 NKSDDKYHFRIRPVLYKVQRQVDHLEHYNVEMQWTTCKRPETYN 46
Db      53 NNASNDLYLYRVQRLIRSORLTTGVREIVTVKIGWTKCRNDTSN 98

```

Search completed: March 18, 2004, 14:30:54  
Job time : 28.8373 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 8.8517 Seconds  
(without alignments)  
499.799 Million cell updates/sec

Title: US-09-941-314-10

Sequence: 1 NKESEDKHYFRIRPVAKVQR.....EYHLNVEMQWTTQCKPPTTN 46

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	43.7	139	2	A45361
2	103	40.6	120	2	S10587
3	103	40.6	127	2	S07085
4	98	38.6	112	1	UDBO
5	95	37.4	140	2	A36163
6	94	37.0	139	1	UDCH
7	94	37.0	146	1	UDHU
8	84	33.1	141	2	B29632
9	81	31.9	162	2	A43428
10	80	31.5	111	2	A28793
11	75	29.5	133	2	UC4556
12	73	28.7	141	1	UDHDP2
13	71	28.0	111	1	UC2040
14	71	28.0	132	2	UC4918
15	71	28.0	141	1	UDHUP1
16	71	28.0	149	2	T31871
17	67	26.4	142	2	A47142
18	63	24.8	438	2	A47702
19	63	24.8	438	2	T52149
20	62	24.4	434	1	KGBO12
21	62	24.4	439	1	KGBOH2
22	61	24.0	436	1	KGBO11
23	61	24.0	621	1	KGBOH1
24	60	23.6	257	2	T03724
25	59	23.2	139	2	B64005
26	59	23.2	139	2	B64005
27	59	23.2	644	1	KGHUI1
28	58	23.0	246	1	H90539
29	58	22.8	448	2	JN0118

30	57.5	22.6	1585	2	T19121	probable protein-t
31	57	22.4	141	2	QJ01470	cystatin S precurs
32	56.5	22.2	325	2	P69784	conserved hypochet
33	56	22.0	200	2	I46051	secreted phosphopr
34	56	22.0	588	2	C95252	L-fucose isomerase
35	56	22.0	588	2	A97117	L-fucose isomerase
36	56	22.0	617	2	S19254	nitrate reductase
37	55.5	21.9	164	2	T31026	hypothetical prote
38	55	21.7	480	2	H86411	protein P1K23.12 (
39	55	21.7	527	2	E90740	probable enzyme [i
40	55	21.7	527	2	G85590	probable enzyme yb
41	55	21.7	527	2	G64818	probable membrane
42	55	21.7	904	1	RDNTNT	nitrate reductase
43	55	21.7	904	1	RDNTNT	nitrate reductase
44	54.5	21.5	1421	2	T49500	hypothetical prote
45	54	21.3	226	2	F83782	transcription regu

## ALIGNMENTS

## RESULT 1

A45361  
cystatin-related epididymal specific protein - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C/Accession: A45361  
R/Cornwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.  
Mol. Endocrinol. 6, 1653-1664, 1992  
A/Title: The CRIS gene: a unique testis-regulated gene related to the cystatin family is  
A/Reference number: A45361; MUID:93078799; PMID:1280328  
A/Accession: A45361  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: nucleic acid  
A/Residues: 1-139 <COR>  
A/Cross-references: GB:549926; NID:9260492; PIDN:AAC53390.1; PID:9260493  
A/Note: sequence extracted from NCBI backbone (NCBIP:118813)  
C/Superfamily: cystatin; cystatin homology  
F;28-139/Domain: cystatin homology <CYS>

Query Match 43.7%; Score 111; DB 2; Length 139;  
Best Local Similarity 42.9%; Pred. No. 2.8e-07;  
Matches 18; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

OY 1 NKESEDKHYFRIRPVAKVQRQVTDHLEHYLNVEMQWTTQCKP 42  
Db 54 NKESEDKHYFVLDKILHAKLQITDRMEYQIDVOISNSCKKP 95

## RESULT 2

S10587  
cystatin C - rat  
C/Species: Rattus sp. (rat)  
C/Date: 21-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 16-Jul-1999  
C/Accession: S10587  
R/Barnard, F.; Barnard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Billiard, M.; Gauch.  
Biochem. Biophys. Res. Commun. 199, 161-166, 1990  
A/Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylation  
A/Reference number: S10587; MUID:90380276; PMID:2400577  
A/Accession: S10587  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-120 <ESN>  
A/Note: 43-Asn was also found  
A/Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala  
F;9-120/Domain: cystatin homology <CYS>

Query Match 40.6%; Score 103; DB 2; Length 120;  
Best Local Similarity 43.2%; Pred. No. 2.7e-06;  
Matches 19; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

OY 1 NKESEDKHYFRIRPVAKVQRQVTDHLEHYLNVEMQWTTQCKP 44

Db 35 NKGSDNAYHSRAIQVVARAKQLVAGINYYLDVEMGRITTCRSQT 78

## RESULT 3

S07085  
cystatin C precursor - rat (fragment)  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 01-Dec-1993 #sequence\_revision 03-Aug-1995 #text\_change 16-Jul-1999  
C/Accession: S07085; S01337; S21109  
R/Collem, T.; Dickson, P.W.; Ennard, F.; Averill, S.; Ribridger, G.P.; Gauthier, F.; Scht  
Eur. J. Biochem. 186, 35-42, 1989  
A/Title: The cDNA structure and expression analysis of the genes for the cysteine protei  
A/Reference number: S07085; MUID:90092122; PMID:2689174  
A/Accession: S07085  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-127 <COL>  
A/Cross-references: EMBL:X16957; NID:956041; PIDN:CAA34831.1; PID:G736290  
R/Ennard, A.; Ennard, F.; Faucher, D.; Gauthier, F.  
FEBS Lett. 236, 475-478, 1988  
A/Title: Two rat homologues of human cystatin C.  
A/Reference number: S01337; MUID:88313020; PMID:3044831  
A/Accession: S01337  
A/Molecule type: protein  
A/Residues: 8-49 <ESN>  
R/Ennard, A.; Ennard, F.; Guillou, F.; Gauthier, F.  
FEBS Lett. 300, 131-135, 1992  
A/Title: Production of the cysteine proteinase inhibitor cystatin C by rat Sertoli cell  
A/Reference number: S21109; MUID:9225121; PMID:1563513  
A/Accession: S21109  
A/Molecule type: protein  
A/Residues: 8, 'XX', 11-20 <ES2>  
C/Superfamily: cystatin: cystatin homology  
C/Keywords: cysteine proteinase inhibitor  
F/16-127/Domain: cystatin homology <CYS>  
F/80-90,104-124/Disulfide bonds: #status predicted

Query Match 40.6%; Score 103; DB 2; Length 127;  
Best Local Similarity 43.2%; Pred. No. 2.8e-06;  
Matches 19; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

Qy 1 NKGSDKYHFRIFRVLKVRQVTDHLEHNLVNMQWTTCCQKPE 44  
Db 42 NKGSDNAYHSRAIQVVARAKQLVAGINYYLDVEMGRITTCRSQT 85

## RESULT 4

UDBO  
cystatin - bovine  
N/Alternate names: thiol proteinase inhibitor  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 28-Feb-1986 #sequence\_revision 28-Feb-1986 #text\_change 06-Dec-1996  
C/Accession: A01271  
R/Hirado, M.; Tanasawa, S.; Sakiyama, F.; Minobe, M.; Fujii, S.  
FEBS Lett. 186, 41-45, 1985  
A/Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase in  
A/Reference number: A01271; MUID:85231205; PMID:3891407  
A/Accession: A01271  
A/Molecule type: protein  
A/Residues: 1-112 <HIR>  
C/Superfamily: cystatin: cystatin homology  
C/Keywords: colostrum; cysteine proteinase inhibitor  
F/2-112/Domain: cystatin homology <CYS>  
F/48-52/Region: inhibitory #status predicted  
F/66-76,90-110/Disulfide bonds: #status predicted

Query Match 38.6%; Score 98; DB 1; Length 112;  
Best Local Similarity 41.9%; Pred. No. 1.1e-05;  
Matches 18; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

Qy 1 NKGSDKYHFRIFRVLKVRQVTDHLEHNLVNMQWTTCCQKPE 43  
Db 42 NKGSDNAYHSRAIQVVARAKQLVAGINYYLDVEMGRITTCRSQT 85

Db 28 NKGSDNAYQSRVYRVARAKQVSGMNYFLDVELGRTTCRSQ 70

## RESULT 5

A36163  
cystatin C precursor - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 16-Jul-1999  
C/Accession: A36163  
R/Solem, M.; Rawson, C.; Lindburg, K.; Barnes, D.  
Biochem. Biophys. Res. Commun. 172, 945-951, 1990  
A/Title: Transforming growth factor beta regulates cystatin C in serum-free mouse embryo  
A/Reference number: A36163; MUID:91054522; PMID:2241983  
A/Accession: A36163  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-140 <SOL>  
A/Cross-references: EMBL:M59470; NID:9192911; PIDN:AAA63298.1; PID:G192912  
C/Superfamily: cystatin: cystatin homology  
F/29-140/Domain: cystatin homology <CYS>  
F/93-103,117-137/Disulfide bonds: #status predicted

Query Match 37.4%; Score 95; DB 2; Length 140;  
Best Local Similarity 40.9%; Pred. No. 3.5e-05;  
Matches 18; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

Qy 1 NKGSDKYHFRIFRVLKVRQVTDHLEHNLVNMQWTTCCQKPE 44  
Db 55 NKGSDNAYHSRAIQVVARAKQLVAGINYYLDVEMGRITTCRSQT 98

## RESULT 6

UDCH

cystatin precursor - chicken  
N/Alternate names: cystatin 1, cysteine proteinase inhibitor; egg-white cystatin  
C/Species: Gallus gallus (chicken)  
C/Date: 03-Aug-1984 #sequence\_revision 12-Apr-1996 #text\_change 29-Oct-1999  
C/Accession: A34456; A01274; S01461; S48159; S04008; JN0789  
R/Colletta, R.; Sakaguchi, Y.; Nagase, H.; Bird, J.W.C.  
J. Biol. Chem. 264, 17164-17169, 1989  
A/Title: Chicken egg white cystatin. Molecular cloning, nucleotide sequence, and tissue c  
A/Reference number: A34456; MUID:9008873; PMID:2793849  
A/Accession: A34456  
A/Molecule type: mRNA  
A/Residues: 1-139 <COL>  
A/Cross-references: GB:J05077; NID:9211714; PIDN:AAA48744.1; PID:G211715  
R/Schwabe, C.; Anastasi, A.; Crow, H.; McDonald, J.K.; Barrett, A.U.  
Biochem. J. 217, 813-817, 1984  
A/Title: Cystatin. Amino acid sequence and possible secondary structure.  
A/Reference number: A01274; MUID:84178305; PMID:6712597  
A/Accession: A01274  
A/Molecule type: protein  
A/Residues: 24-139 <SCH>  
R/Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983  
A/Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin  
A/Reference number: S01461; MUID:84110059; PMID:6662498  
A/Accession: S01461  
A/Molecule type: protein  
A/Residues: 24-139 <TUR>  
R/Anastasi, A.; Brown, M.A.; Kemhavi, A.A.; Nicklin, M.U.H.; Sayers, C.A.; Sunter, D.C.,  
Biochem. J. 211, 129-138, 1983  
A/Title: Cystatin, a protein inhibitor of cysteine proteinases. Improved purification fr  
A/Reference number: A37514; MUID:83256421; PMID:6490805  
A/Contents: annotation; characterization of protein  
R/Grubb, A.; Lotberg, H.; Barrett, A.U.  
FEBS Lett. 170, 370-374, 1984  
A/Title: The disulfide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
A/Reference number: S01462  
A/Contents: annotation; disulfide bonds  
R/Auerwald, E.A.; Naegler, D.K.; Schulze, A.U.; Engn, R.A.; Genenger, G.; Machleidt, W.,  
Eur. J. Biochem. 224, 407-415, 1994  
A/Title: Production, inhibitory activity, folding and conformational analysis of an N-ter

A/Reference number: S48159; MUID:95010016; PMID:7925354  
 A/Accession: S48159  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 24-139 <AUB>  
 R/Lieber, B.; Krieglstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.  
 FEBS Lett. 249, 162-168, 1989  
 A/Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.  
 A/Reference number: S04008; MUID:89252033; PMID:2721673  
 A/Accession: S04008  
 A/Molecule type: protein  
 A/Residues: 97-114 <LAB>  
 R/Colletta, R.; Bird, J.W.C.  
 Gene 130, 175-181, 1993  
 A/Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping  
 A/Reference number: JN0789; MUID:93366172; PMID:8359684  
 A/Accession: JN0789  
 A/Molecule type: DNA  
 A/Residues: 1-139 <CO2>  
 A/Cross-references: GB:M95725  
 A/Note: authors failed to translate the codon for residue 115-Tyr  
 C/Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases  
 C/Genetics:  
 A/Genes: Can  
 A/Introns: 76/3; 114/3  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein  
 F/1-23/Domain: signal sequence #status predicted <SIG>  
 F/24-139/Product: cystatin, long form #status experimental <CYLF>  
 F/30-139/Domain: cystatin homology <CYS>  
 F/32-139/Product: cystatin, short form #status experimental <CYSB>  
 F/76-80/Region: inhibitory #status predicted  
 F/94-104,118-138/Diethylidide bonds: #status experimental  
 F/103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 37.0%; Score 94; DB 1; Length 139;  
 Best Local Similarity 43.9%; Pred. No. 4.8e-05;  
 Matches 18; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKESDDKYPFRIFRVLKVRQVTDHLEHYNEMQMTTCOK 41  
 Db 56 NRASNDKYSRRVVRVIAKQQLVSGIKYILQVEIGRTTCRK 96

RESULT 7  
 UDHU  
 Cystatin C precursor [validated] - human  
 N/Alternate names: gamma-CSP; gamma-trace; neuroendocrine basic polypeptide; post-gamma  
 C/Species: Homo sapiens (man)  
 C/Date: 06-Jul-1982 #sequence revision 31-Mar-1991 #text change 08-Dec-2000  
 C/Accession: S10216; S00004; J10095; A33400; S02751; A01270; A25434; S12288; A32732; A60  
 R/Abrahamson, M.; Olafsson, I.; Palasdottrir, A.; Ulvbaeck, M.; Lundvall, A.; Jansson, O.  
 Biochem. J. 268, 287-294, 1990  
 A/Title: Structure and expression of the human cystatin C gene.  
 A/Reference number: S10216; MUID:90303202; PMID:2363674  
 A/Accession: S10216  
 A/Molecule type: DNA  
 A/Residues: 1-146 <AB1>  
 A/Cross-references: EMBL:X52255; NID:930257; PIDN:CAA3497.1; PID:9296643  
 R/Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundvall, A.  
 FEBS Lett. 216, 229-233, 1987  
 A/Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the  
 A/Reference number: S00004; MUID:87219149; PMID:3495457  
 A/Accession: S00004  
 A/Molecule type: mRNA  
 A/Residues: 1-146 <AB2>  
 A/Cross-references: EMBL:X05607; NID:930371; PIDN:CAA2096.1; PID:9755738  
 R/Levy, E.; Lopez-Otin, C.; Ghiso, J.; Gellner, D.; Frangione, B.  
 J. Exp. Med. 169, 1771-1778, 1989  
 A/Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a  
 A/Reference number: J10095; MUID:89235594; PMID:2541223  
 A/Accession: J10095  
 A/Molecule type: DNA

A/Residues: 1-146 <LEV>  
 A/Cross-references: GB:X61681; NID:930367; PIDN:CAA43856.2; PID:94490944  
 A/Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit  
 e)  
 R/Satch, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada, H.  
 Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989  
 A/Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which  
 A/Reference number: A33400; MUID:89350949; PMID:2764935.  
 A/Accession: A33400  
 A/Molecule type: DNA  
 A/Residues: 1-24, 'T', '26-146 <SA1>  
 A/Cross-references: GB:M27889; GB:M27891; NID:9181385; PIDN:AA52164.1; PID:921  
 R/Ghiso, J.; Cowan, N.; Frangione, B.  
 Biol. Chem. Hoppe-Seyler 369, 205-208, 1988  
 A/Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron  
 A/Reference number: S02751; MUID:89076507; PMID:3264504  
 A/Accession: S02751  
 A/Molecule type: DNA  
 A/Residues: 82-119 <GH2>  
 A/Cross-references: EMBL:M2769  
 A/Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown  
 R/Grubb, A.; Lofberg, H.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982  
 A/Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in the  
 A/Reference number: A01270; MUID:82222268; PMID:6283552  
 A/Accession: A01270  
 A/Molecule type: protein  
 A/Residues: 27-131, 'S', '133-146 <GRU>  
 R/Ghiso, J.; Jansson, O.; Frangione, B.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986  
 A/Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland ty  
 A/Reference number: A25434; MUID:86206076; PMID:3517880  
 A/Accession: A25434  
 A/Molecule type: protein  
 A/Residues: 37-93, 'Q', '95-146 <GH1>  
 R/Turk, V.; Brzin, J.; Loner, M.; Ritonja, A.; Brodtkin, M.; Borchardt, U.; Machleidt, W.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983  
 A/Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatir  
 A/Reference number: S01461; MUID:84110053; PMID:6662458  
 A/Accession: S12288  
 A/Molecule type: protein  
 A/Residues: 27-73 <TUR>  
 R/Brzin, J.; Popovic, T.; Turk, V.  
 Biochem. Biophys. Res. Commun. 118, 103-109, 1984  
 A/Title: Human cystatin, a new protein inhibitor of cysteine proteinases.  
 A/Reference number: A32732; MUID:84128015; PMID:6365094  
 A/Accession: A32732  
 A/Molecule type: protein  
 A/Residues: 27-76 <BR2>  
 R/Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.  
 Scand. J. Clin. Lab. Invest. 50, 85-93, 1990  
 A/Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst  
 A/Reference number: A60552; MUID:90193615; PMID:2315647  
 A/Accession: A60552  
 A/Molecule type: protein  
 A/Residues: 27-49, 'XX', '52-64 <OLA>  
 A/Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d  
 e defective gene is not present in CSF but is found instead in amyloid deposits  
 R/Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.  
 Biol. Chem. Hoppe-Seyler 371, 575-580, 1990  
 A/Title: Different forms of human cystatin C.  
 A/Reference number: S10607; MUID:91025625; PMID:2222856  
 A/Accession: S10607  
 A/Molecule type: protein  
 A/Residues: 27-53 <POP>  
 A/Experimental source: urine, kidney disease  
 A/Note: truncated forms with amino ends at positions 35 and 36 of the precursor were also  
 R/Grubb, A.; Lofberg, H.; Barrett, A.J.  
 FEBS Lett. 170, 370-374, 1984  
 A/Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
 A/Reference number: S01462  
 A/Contents: annotation; disulfide bonds  
 R/Berti, P.J.; Storer, A.C.

Biochem. J. 302, 411-416, 1994  
 A/Title: Local pH-dependent conformational changes leading to proteolytic susceptibility  
 A/Reference number: S55305; MUID:94379969; PMID:8092991  
 A/Accession: S55305  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 27-49/106-146 <BER>  
 C/Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fluid patients with certain autoimmune diseases.  
 C/Comment: This protein is an inhibitor of cysteine proteinases and may serve an important function in the regulation of the immune response.  
 C/Comment: A mutant cystatin C, with 94-qln, is deposited in hereditary cerebral hemorrhage with amyloidosis (HCHWA-C).  
 A/Accession: GDB:CST3  
 A/Cross-references: GDB:119817; OMIM:105150  
 A/Map position: 20p11.2-20p11.2  
 A/Intons: 81/3; 119/3  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline  
 F/1-26/Domain: signal sequence #status predicted <SIG>  
 F/27-146/Product: cystatin C #status experimental <MAT>  
 F/35-146/Domain: cystatin homology <CYS>  
 F/81-85/Region: inhibitory #status predicted  
 F/82/Modified site: hydroxyproline (Pro) (partial) #status experimental  
 F/99-109,123-143/Disulfide bonds: #status experimental

Query Match 37.0%; Score 94; DB 1; Length 146;  
 Best Local Similarity 39.5%; Pred. No. 5e-05;  
 Matches 17; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

1 NKESDDKXHFRIFRVLKVGQVTDHLEHLNVEWMTTCCK 43  
 Db 61 NKASNDWYHSRALQVVARQIVGVNVPFLDVEIGRTCTKTQ 103

RESULT 8  
 Cystatin SA precursor - human  
 C/Species: Homo sapiens (man)  
 C/Date: 31-Mar-1989 #sequence revision 30-Jun-1989 #text\_change 16-Jul-1999  
 C/Accession: B29632; S02490; A41422; B27015  
 R/Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.  
 Gene 61, 329-338, 1987  
 A/Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three men  
 A/Reference number: A91589; MUID:88185836; PMID:3446578  
 A/Accession: B29632  
 A/Molecule type: DNA  
 A/Residues: 1-141 <SAI>  
 A/Cross-references: GB:M19673; GB:M19170; NID:g186403; PIDN:AAA3116.1; PID:g386826  
 A/Note: the authors translated the codon GAC for residue 129 as Asn  
 R/Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.  
 Biol. Chem. Hoppe-Seyler 369, 191-197, 1988  
 A/Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily  
 A/Reference number: S02489; MUID:89076505; PMID:33202964  
 A/Accession: S02490  
 A/Status: not compared with conceptual translation  
 A/Molecule type: DNA  
 A/Residues: 21-141 <SA2>  
 R/Isemura, S.; Saitoh, E.; Sanada, K.  
 J. Biochem. 102, 693-704, 1987  
 A/Title: Characterization and amino acid sequence of a new acidic cysteine proteinase in  
 A/Reference number: A41422; MUID:88139220; PMID:3436950  
 A/Accession: A41422  
 A/Molecule type: protein  
 A/Residues: 25-141 <ISE>  
 R/Isemura, S.; Saitoh, E.; Sanada, K.; Isemura, M.; Ito, S.  
 in Cysteine Proteinases and Their Inhibitors, Turk, V., ed., pp.497-505, Walter de Gruyter  
 A/Title: Cystatin S and the related cysteine proteinase inhibitors in human saliva.  
 A/Reference number: A27015  
 A/Accession: B27015  
 A/Molecule type: protein  
 A/Residues: 25-134, 'D', 136-141 <IS2>  
 C/Genetics: GDB:CST2

A/Cross-references: GDB:119816; OMIM:123856  
 A/Map position: 20p11.2-20p11.2  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 27-49/106-146 <BER>  
 C/Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fluid patients with certain autoimmune diseases.  
 C/Comment: This protein is an inhibitor of cysteine proteinases and may serve an important function in the regulation of the immune response.  
 C/Comment: A mutant cystatin C, with 94-qln, is deposited in hereditary cerebral hemorrhage with amyloidosis (HCHWA-C).  
 A/Accession: GDB:CST3  
 A/Cross-references: GDB:119817; OMIM:105150  
 A/Map position: 20p11.2-20p11.2  
 A/Intons: 81/3; 119/3  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline  
 F/1-26/Domain: signal sequence #status predicted <SIG>  
 F/27-146/Product: cystatin C #status experimental <MAT>  
 F/35-146/Domain: cystatin homology <CYS>  
 F/81-85/Region: inhibitory #status predicted  
 F/82/Modified site: hydroxyproline (Pro) (partial) #status experimental  
 F/99-109,123-143/Disulfide bonds: #status experimental

Query Match 33.1%; Score 84; DB 2; Length 141;  
 Best Local Similarity 32.6%; Pred. No. 0.00099;  
 Matches 14; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

1 NKESDDKXHFRIFRVLKVGQVTDHLEHLNVEWMTTCCK 43  
 Db 56 NKATEDRYRLRLRLVARQIVGVNVPFLDVEIGRTCTKTQ 98

RESULT 9  
 onchocystatin - nematode (Onchocerca volvulus)  
 N/Alternate names: cysteine proteinase inhibitor; onchocerciasis antigen  
 C/Species: Onchocerca volvulus  
 C/Date: 04-Mar-1993 #sequence revision 18-Nov-1994 #text\_change 17-Mar-2000  
 C/Accession: A43428; B43927  
 R/Lustigman, S.; Brozman, B.; Hulma, T.; Prince, A.M.; McKerrow, J.H.  
 J. Biol. Chem. 267, 17339-17346, 1992  
 A/Title: Molecular cloning and characterization of onchocystatin, a cysteine proteinase  
 A/Reference number: A43428; MUID:92381053; PMID:1512269  
 A/Accession: A43428  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-162 <LUS>  
 A/Cross-references: GB:M37105; NID:g159905; PID:g159906  
 A/Note: Sequence extracted from NCBI backbone (NCBI:111962, NCBI:111963)  
 R/Chandrasekhar, R.; Masood, K.; Alvarez, R.M.; Oguntirade, A.F.; Lujan, R.; Richards Jr  
 J. Clin. Invest. 88, 1460-1466, 1991  
 A/Title: Molecular cloning and characterization of recombinant parasite antigens for imm  
 A/Reference number: A43927; MUID:92042729; PMID:1840605  
 A/Accession: B43927  
 A/Molecule type: mRNA  
 A/Residues: 'P', 37-57, 'A', 59-71, 'R', 73-83, 'N', 85-126, 'W', 128-162 <CHA>  
 A/Cross-references: GB:M60279; NID:g159888  
 A/Note: sequence extracted from NCBI backbone (NCBI:65111, NCBI:65113)  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: cysteine proteinase inhibitor

Query Match 31.9%; Score 81; DB 2; Length 162;  
 Best Local Similarity 34.1%; Pred. No. 0.0029;  
 Matches 14; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

1 NKESDDKXHFRIFRVLKVGQVTDHLEHLNVEWMTTCCK 41  
 Db 77 NEQSDNDEYHMPKILKVSQVAVGVKMDVQVARSQCKK 117

RESULT 10  
 Cystatin - puff adder  
 C/Species: Bitis arietans (puff adder)  
 C/Date: 15-Dec-1988 #sequence revision 15-Dec-1988 #text\_change 30-Sep-1993  
 C/Accession: A28793  
 R/Ritonga, A.; Evans, H.J.; Machleidt, W.; Barrett, A.J.  
 Biochem. J. 246, 799-802, 1987  
 A/Title: Amino acid sequence of a cystatin from venom of the African puff adder (Bitis a  
 A/Reference number: A28793; MUID:88076861; PMID:3500714  
 A/Accession: A28793  
 A/Molecule type: protein  
 A/Residues: 1-111 <RIT>  
 C/Superfamily: cystatin; cystatin homology

Query Match 31.5%; Score 80; DB 2; Length 111;  
 Best Local Similarity 39.0%; Pred. No. 0.0025;  
 Matches 16; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

1 NKESDDKXHFRIFRVLKVGQVTDHLEHLNVEWMTTCCK 41

Db 27 NAGSKNDYFKRRVVAQSGVSVGVKXYLMELKTTCK 67

## RESULT 11

JC4536  
Cysteatin precursor - horseshoe crab (Trachyleus tridentatus)  
N/Alternate names: cystatin L; cysteine proteinase inhibitor  
C/Species: Trachyleus tridentatus  
C/Date: 15-Feb-1996 #sequence\_revision 19-Apr-1996 #text\_change 03-Dec-1999  
C/Accession: JC4536; PC4122  
R/Author: K.L.; Kawabata, S.; Hirata, M.; Miyagi, M.; Tanasawa, S.; Iwanaga, S.  
J. Biochem. 119, 85-94, 1996  
A/Title: A cysteine protease inhibitor scored in the large granules of horseshoe crab he  
A/Reference number: JC4536; MUID:97063112; PMID:8907180  
A/Accession: JC4536  
A/Molecule type: mRNA  
A/Residues: 1-133 <AGA>  
A/Accession: PC4122  
A/Molecule type: protein  
A/Residues: 26-57/60-105/107-113/115-128 <AG2>  
A/Experimental source: hemocytes  
C/Comment: This protein belongs to the cystatin family and it is a single-chain protein  
activity against Gram-negative bacteria, defense against invading microbes, and response t  
C/Superfamily: cystatin; cystatin homology  
C/Keywords: cysteine proteinase inhibitor; hemolymph; pyroglutamic acid  
F/1-19/Domain: signal sequence #status predicted <SIG>  
F/20-133/Product: limulus-cystatin #status predicted <MAT>  
F/21-132/Domain: cystatin homology <CYS>  
F/20/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F/85-98,109-129/Disulfide bonds: #status predicted

Query Match 29.5%; Score 75; DB 2; Length 133;  
Best Local Similarity 30.2%; Pred. No. 0.014;  
Matches 13; Conservative 11; Mismatches 19; Indels 0; Gaps 0;  
1 NKESDDKYHFRIRVLRVQVTDHLEHLNVEQMWTTCCK 43  
47 SRSNSLYHKLKTKHARTQVSGINVEVETGTTCKSE 89

## RESULT 12

UDHUP2  
Cystatin SN precursor (validated) - human  
N/Alternate names: cystatin SA-I  
C/Species: Homo sapiens (man)  
C/Date: 28-May-1986 #sequence\_revision 08-Feb-1996 #text\_change 08-Dec-2000  
C/Accession: A28110; S02489; A29632; A01273; S19279  
R/Author: Heshimi, I.; Dickinson, D.P.; Levine, M.J.  
J. Biol. Chem. 263, 9381-9387, 1988  
A/Title: Purification, molecular cloning, and sequencing of salivary cystatin, SA-I.  
A/Reference number: A28110; MUID:88243825; PMID:2837486  
A/Accession: A28110  
A/Molecule type: mRNA  
A/Residues: 1-141 <LH>  
A/Cross-references: GB:03870; NID:g337751; PID:AAA60299.1; PID:g337752  
R/Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.  
Biol. Chem. Hoppe-Seyler 369, 191-197, 1988  
A/Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily  
A/Reference number: S02489; MUID:89076505; PMID:3202964  
A/Accession: S02489  
A/Status: not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 21-141 <SA2>  
R/Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.  
Gene 61, 329-338, 1987  
A/Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three mem  
A/Reference number: A91589; MUID:88168635; PMID:3446578  
A/Accession: A29632  
A/Molecule type: DNA  
A/Residues: 1-86, '1', 88-141 <SA1>  
R/Isemura, S.; Saitoh, E.; Sanada, K.  
FEBS Lett. 198, 145-149, 1986

A/Title: Characterization of a new cysteine proteinase inhibitor of human saliva, cystati  
A/Reference number: A01273; MUID:86164938; PMID:3514222

A/Accession: A01273  
A/Molecule type: protein  
A/Residues: 29-141 <ISE>  
R/Ramasubbu, N.; Reddy, M.S.; Bergey, E.J.; Harasathy, G.G.; Soni, S.D.; Levine, M.J.  
Biochem. J. 280, 341-352, 1991  
A/Title: Large-scale purification and characterization of the major phosphoproteins and  
A/Reference number: S19279; MUID:92082465; PMID:11747107  
A/Accession: S19279  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 21-55 <RAM>  
C/Comment: Human saliva appears to contain several cysteine proteinase inhibitors that a  
ences. Cystatin SN, with a pI of 7.5, is a much better inhibitor of papain and dipeptidyl  
C/Genetics:  
A/Gene: GDB:CS71  
A/Cross-references: GDB:119815; OMIM:123855  
A/Map position: 20p11.2-20p11.2  
C/Superfamily: cystatin; cystatin homology  
C/Keywords: cysteine proteinase inhibitor; extracellular protein; saliva  
F/1-20/Domain: signal sequence #status predicted <SIG>  
F/21-141/Product: cystatin SA-I #status experimental <MAT1>  
F/29-141/Product: cystatin SN #status experimental <MAT2>  
F/30-141/Domain: cystatin homology <CYS>  
F/76-80/Region: inhibitory #status predicted  
F/94-104,118-138/Disulfide bonds: #status predicted

Query Match 28.7%; Score 73; DB 1; Length 141;  
Best Local Similarity 34.9%; Pred. No. 0.027;  
Matches 15; Conservative 9; Mismatches 19; Indels 0; Gaps 0;  
1 NKESDDKYHFRIRVLRVQVTDHLEHLNVEQMWTTCCK 43  
Db 56 NKATKDYRRPLRLARQQVGVVFPDVEVGRITCKSQ 98

## RESULT 13

JC2040  
Cystatin - chum salmon  
N/Alternate names: cysteine proteinase inhibitor  
C/Species: Oncorhynchus keta (chum salmon)  
C/Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 16-Jul-1999  
C/Accession: JC2040  
R/Koide, Y.; Noso, T.  
Biochim. Biotechnol. Biochem. 58, 164-169, 1994  
A/Title: The complete amino acid sequence of pituitary cystatin from chum salmon.  
A/Reference number: JC2040; MUID:94162738; PMID:7764512  
A/Accession: JC2040  
A/Molecule type: protein  
A/Residues: 1-111 <KOI>  
C/Comment: The intracellular role of this protein is the inhibition of intralysosomal pr  
C/Superfamily: cystatin; cystatin homology  
C/Keywords: cysteine proteinase inhibitor  
F/2-111/Domain: cystatin homology <CYS>  
F/48-52/Region: inhibitory  
F/89-109/Disulfide bonds: #status experimental

Query Match 28.0%; Score 71; DB 1; Length 111;  
Best Local Similarity 31.7%; Pred. No. 0.039;  
Matches 13; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

1 NKESDDKYHFRIRVLRVQVTDHLEHLNVEQMWTTCCK 41  
Db 28 NKXTNDFVROQVAKVNAQKQVSGMKYFTVQGRTPCRK 68

## RESULT 14

JC4918  
Cystatin precursor - chum salmon  
C/Species: Oncorhynchus keta (chum salmon)  
C/Date: 26-Sep-1996 #sequence\_revision 01-Nov-1996 #text\_change 21-Jan-2000  
C/Accession: JC4918





CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; AL096677; CAC13170.1; -  
DR EMBL; AL096677; CAC17423.1; -  
DR HSSP; P01038; 1A90.  
DR Genew; HGNC:15959; CST11.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00031; cystatin.1.  
DR SMART; SM00043; CY; 1.  
DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
KM Thiol protease inhibitor; Signal; Alternative splicing.  
FT SIGNAL 1 25  
FT CHAIN 26 137  
FT SITE 75 79  
FT DISULFID 93 101  
FT CARBOHYD 114 134  
FT VARSPIC 131 131  
FT 76 110  
SQ SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;  
Query Match 100.0%; Score 254; DB 1; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.6e-26;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIPLVLYKQROVTDHLEYHLNVEMQWTTCKPPTN 46  
Db 55 NKESDDKXHFRIPLVLYKQROVTDHLEYHLNVEMQWTTCKPPTN 100

## RESULT 2

CST1\_MOUSE  
ID CST1\_MOUSE STANDARD; PRT; 139 AA.

AC Q9D269;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cystatin 11 precursor.  
GN CST11.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Epididymis;  
RX MEDLINE=21085660; PubMed=11217851;

RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kaubuka T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann M., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nakado T., Pesole G., Quackenbush J.,  
RA Schirral L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamitani M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombereis P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -1- SIMILARITY: Belongs to the cystatin family.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; AK020300; BAB32061.1; -  
DR HSSP; P01034; 1G96.  
DR MGD; MGI:1925490; Cst11.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00031; cystatin.1.  
DR SMART; SM00043; CY; 1.  
DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
KM Thiol protease inhibitor; Signal.  
FT SIGNAL 1 28  
FT CHAIN 29 139  
FT SITE 76 80  
FT DISULFID 94 102  
FT CARBOHYD 115 135  
FT 134 134  
SQ SEQUENCE 139 AA; 16217 MW; F228D9815FA32640 CRC64;  
Query Match 63.8%; Score 162; DB 1; Length 139;  
Best Local Similarity 64.4%; Pred. No. 2.2e-14;  
Matches 29; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIPLVLYKQROVTDHLEYHLNVEMQWTTCKPPTN 45  
Db 56 NKESDDKXHFRIPLVLYKQROVTDHLEYHLNVEMQWTTCKPPTN 100

## RESULT 3

CST8\_HUMAN  
ID CST8\_HUMAN STANDARD; PRT; 142 AA.

AC O60676;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin  
DE 8).  
GN CST8 OR CREB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=95344753; PubMed=7619504;

RA Cornwall G.A., Hann S.R.,  
RT "Transient appearance of CREB protein during spermatogenesis and  
RT caput epididymal sperm maturation."  
RL Mol. Reprod. Dev. 41:37-46(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blake S.E., Bridgman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Cosby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dham P.D., Dunn M.,  
RA Ellington D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Grafham D., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Humble E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehtvaeslahti M.H., Leverna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McInay K., McMurtry A.A.,



DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin  
 b)  
 GN CST8 OR CRES.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 ON NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Epididymis;  
 RX MEDLINE=99247899; PubMed=10229662;  
 RA Cornwall G.A., Hsia N., Sutton H.G.;  
 RT "Structure, alternative splicing and chromosomal localization of the  
 RT cystatin-related epididymal spermatogenic gene.";  
 RL Biochem. J. 340:85-93(1999).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF090692; AAC36317.1; -.  
 CC DR HSSP: P01034; 1G96.  
 CC DR InterPro: IPR00010; Cystatin.  
 CC DR Pfam: PF00031; Cystatin; 1.  
 CC DR SMART: SM00043; Cy:1.  
 CC KM Thiol protease inhibitor; Signal.  
 CC FT SIGNAL 1 19 POTENTIAL.  
 CC FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 CC FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).  
 CC FT DISULFID 95 105 BY SIMILARITY.  
 CC FT DISULFID 119 139 BY SIMILARITY.  
 CC FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SQ SEQUENCE 142 AA; 16246 MW; FB873FMA6B6CAB34 CRC64;  
 CC  
 CC Query Match 41.7%; Score 106; DB 1; Length 142;  
 CC Best Local Similarity 42.9%; Pred. No. 5.3e-07; Indels 0; Gaps 0;  
 CC Matches 18; Conservative 11; Mismatches 13;  
 CC  
 CC QY 1 NKESDDKYHFRIPVLYKVRQVTDHLEHNLNEMQWTTCKRP 42  
 CC Db 57 NKGSDDKYLFLDHTLHATLQITRMEYHIDVQISRSNCRKP 98  
 CC  
 CC RESULT 6  
 CC CYTC RAT STANDARD; PRT; 127 AA.  
 CC ID CYTC RAT  
 CC AC P14841; 01-APR-1990 (Rel. 14, Created)  
 CC DT 01-APR-1990 (Rel. 14, Last sequence update)  
 CC DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Cystatin C precursor (Fragment).  
 CC GN CST3.  
 CC OS Rattus norvegicus (Rat).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CC ON NCBI\_TaxID=10116;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=Buffalo;  
 CC RX MEDLINE=90092122; PubMed=2689174;  
 CC RA Cole T., Dickson P.W., Esmad F., Averill F., Risbridger G.,  
 CC Gauthier F., Schreiber G.;  
 CC RT "The CDNA structure and expression analysis of the genes for the

RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin  
 RT in rat brain.";  
 RL Eur. J. Biochem. 186:35-42(1989).  
 RN [2]  
 RN SEQUENCE OF 8-127.  
 RX MEDLINE=90380276; PubMed=2400577;  
 RA Esmad F., Esmad A., Faucher D., Capony J.-P., Derancourt J.,  
 RA Briard M., Gauthier F.;  
 RT "Rat cystatin C: the complete amino acid sequence reveals a site for  
 RT N-glycosylation.";  
 RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).  
 RN [3]  
 RN SEQUENCE OF 8-49.  
 RX MEDLINE=88313020; PubMed=3044831;  
 RA Esmad A., Esmad F., Faucher D., Gauthier F.;  
 RT "Two rat homologues of human cystatin C.";  
 RL FEBS Lett. 236:475-478(1988).  
 RN [4]  
 RN SEQUENCE OF 8-20.  
 RC TISSUE=Sertoli cells;  
 RX MEDLINE=9225121; PubMed=1563513;  
 RA Esmad A., Esmad F., Guillo F., Gauthier F.;  
 RT "Production of the cysteine proteinase inhibitor cystatin C by rat  
 RT Sertoli cells.";  
 RL FEBS Lett. 300:131-135(1992).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 CC thought to serve an important physiological role as a local  
 CC regulator of this enzyme activity. Known to inhibit cathepsin B,  
 CC H. and L.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X16957; CA334831.1; -.  
 CC DR PIR: S07085; S07085.  
 CC DR PIR: S10587; S10587.  
 CC DR HSSP: P01034; 1G96.  
 CC DR InterPro: IPR00010; Cystatin.  
 CC DR Pfam: PF00031; Cystatin; 1.  
 CC DR SMART: SM00043; Cy:1.  
 CC DR PROSITE: PS00287; CYSTATIN; 1.  
 CC KM Thiol protease inhibitor; Signal.  
 CC FT NON TER 1 1  
 CC FT CHAIN 1 7  
 CC FT SIGNAL <1 7  
 CC FT ACT\_SITE 18 127 CYSTATIN C.  
 CC FT SITE 62 66 REACTIVE SITE.  
 CC FT DISULFID 80 90 SECONDARY AREA OF CONTACT.  
 CC FT DISULFID 104 124 BY SIMILARITY.  
 CC FT CONFLICT 25 25 A -> E (IN REF. 2).  
 CC SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;  
 CC  
 CC Query Match 40.6%; Score 103; DB 1; Length 127;  
 CC Best Local Similarity 43.2%; Pred. No. 1.2e-06;  
 CC Matches 19; Conservative 11; Mismatches 14; Indels 0; Gaps 0;  
 CC  
 CC QY 1 NKESDDKYHFRIPVLYKVRQVTDHLEHNLNEMQWTTCKRP 44  
 CC Db 42 NKGSNDYHSPRAIQVARRQLVAGINYLVDVEMGRITTKRSQT 85  
 CC  
 CC RESULT 7  
 CC CYTC RABIT STANDARD; PRT; 148 AA.  
 CC ID CYTC RABIT  
 CC AC O97862; 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin C precursor.  
 GN CST3.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 RN NCBI\_TaxID=9986;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Japanese white; TISSUE=Bone;  
 RX MEDLINE=98424349; PubMed=9753427;  
 RA Kobori M., Ikeda Y., Nara H., Kamegawa M., Nojima H.,  
 RA Kawachi H.;  
 RT "Large scale isolation of osteoclast-specific genes by an improved  
 RT method involving the preparation of a subtracted cDNA library.";  
 RL Genes Cells 3:459-475(1998).  
 CC -1- FUNCTION: This is a thiol proteinase inhibitor.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB009342; BAA75921.1; -  
 CC HSSP; P01034; 1G96.  
 CC InterPro: IPR000010; Cystatin.  
 CC Pfam: PF00031; Cystatin; 1.  
 DR SMART; SM00043; Cy 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
 DR Thiol protease inhibitor; Signal.  
 KW SIGNAL 1 28  
 FT CHAIN 29 148 POTENTIAL.  
 FT ACT SITE 39 39 REACTIVE SITE.  
 FT SITE 83 87 SECONDARY AREA OF CONTACT.  
 FT DISULFID 101 111 BY SIMILARITY.  
 FT DISULFID 125 145  
 SQ SEQUENCE 148 AA; 16346 MW; 1523C8311695B9A CRC64;  
 Query Match 39.84; Score 101; DB 1; Length 148;  
 Best Local Similarity 40.94; Pred. No. 2.5e-06;  
 Matches 18; Conservative 13; Mismatches 13; Indels 0; Gaps 0;  
 QY 1 NKESDDKYHPRIPVLYKQKQVTDHLEHYNVEMQWTTCKQKPT 44  
 DB 63 NKSDNRHSHRALQVVARARQIVSGVYVYDLIGRTTCKTQY 106  
 RESULT 8  
 CYTC MOUSE STANDARD; PRT; 140 AA.  
 ID CYTC\_MOUSE  
 AC P214EO;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin C precursor (Cystatin 3).  
 GN CST3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C; TISSUE=Brain;  
 RX MEDLINE=91054522; PubMed=2241893;  
 RA Solem M., Rawson C., Lindburg K., Barnes D.;  
 RT "Transforming growth factor beta regulates cystatin C in serum-free  
 RT mouse embryo (SPME) cells.";  
 RL Biochem. Biophys. Res. Commun. 172:945-951(1990).  
 [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv; TISSUE=Liver;  
 RX MEDLINE=95137392; PubMed=7835704;  
 RA Huh C., Nagle J.W., Kozak C.A., Abrahamson M., Karlsson S.;  
 RT "Structural organization, expression and chromosomal mapping of the  
 RT mouse cystatin-C-encoding gene (Cst3).";  
 RL Gene 152:221-226(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ILS, and ISB;  
 RX MEDLINE=21363810; PubMed=11471062;  
 RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,  
 RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;  
 RT "High-throughput sequence identification of gene coding variants  
 RT within alcohol-related OTUs.";  
 RL Mamm. Genome 12:657-663(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Bha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Ketterman M., Madan A.C., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,  
 RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 CC thought to serve an important physiological role as a local  
 CC regulator of this enzyme activity.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M59470; AAA63298.1; -  
 CC EMBL; U10098; AAB41056.1; -  
 CC EMBL; AF483486; AAL90760.1; -  
 CC EMBL; AF483487; AAL90761.1; -  
 CC EMBL; BC020772; AAH02072.1; -  
 CC PIR; A36163; A36163.  
 DR HSSP; P01034; 1G96.  
 DR MGD; MGI:102519; Cst3.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 DR SMART; SM00043; Cy 1.  
 KW Thiol protease inhibitor; Signal.  
 KW SIGNAL 1 20  
 FT CHAIN 21 140  
 FT ACT SITE 31 31 CYSTATIN C.  
 FT SITE 75 79 REACTIVE SITE.  
 FT DISULFID 93 103 SECONDARY AREA OF CONTACT.  
 FT DISULFID 117 137 BY SIMILARITY.  
 FT CONFLICT 16 16 BY SIMILARITY.  
 FT CONFLICT 84 84 A -> G (IN REF. 1).  
 FT CONFLICT 84 84 L -> F (IN REF. 1).

SQ SEQUENCE 140 AA; 15531 MW; 3A563406DD58D0F5 CRC64;

Query Match 39.0%; Score 99; DB 1; Length 140;  
Best Local Similarity 43.2%; Pred. No. 4.4e-06;  
Matches 19; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKESDKYHFRIFRVLKVQRQVTDHLEHYNEMQWTTCKR 44  
DB 55 NKASNDAYHSRAIOVVARAKQIVAGVNYFLDVEGMRTTCKRSQ 98

## RESULT 9

CYT\_C\_BOVIN STANDARD; PRT; 148 AA.

AC P01035;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1998 (Rel. 38, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin C precursor (Colostrum thiol proteinase inhibitor).  
GN CST3.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
NCBI\_TaxID=9913;  
[1]  
SEQUENCE FROM N.A., SEQUENCE OF 66-83, AND CHARACTERIZATION.  
RC TISSUE=Cerebrospinal fluid, and Choroid plexus;  
RX MEDLINE=86094199; PubMed=9434110;  
RA Oleson S.-L., Ek B., Wilm M., Broberg S., Rask L., Bjork I.,  
RT "Molecular cloning and N-terminal analysis of bovine cystatin C  
RT identification of a full-length N-terminal region."  
RL Biochim. Biophys. Acta 1343:203-210(1997).  
[2]  
RP SEQUENCE OF 37-148.  
RX MEDLINE=85231205; PubMed=3891407;  
RA Hirado M., Tsunawake S., Sakiyama F., Ninobe M., Fujii S.,  
RT "Complete amino acid sequence of bovine colostrum low-Mr cysteine  
RT proteinase inhibitor."  
RL FEBS Lett. 186:41-45(1985).  
CC -1- FUNCTION: This is a thiol proteinase inhibitor.  
CC -1- MASS SPECTROMETRY: MW=13420, METHOD=MALDI.  
CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Y10811; CAA71771.1; -.  
DR HSSP: P01034; I966.  
DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; cystatin.1.  
DR SMART: SM00043; CY.1.  
DR PROSITE: PS00287; CYSTATIN.1.  
KW Thiol protease inhibitor; Signal; Pyroglutamate carboxylic acid.  
FT SIGNAL 1 30 PROBABLE.  
FT CHAIN 1 148 CYSTATIN C.  
FT MOD\_RES 31 31 PYROGLUTAMATE CARBOXYLIC ACID (PROBABLE).  
FT ACT\_SITE 40 40 REACTIVE SITE.  
FT SITE 84 88 SECONDARY AREA OF CONTACT.  
FT DISULFID 102 112 BY SIMILARITY.  
FT DISULFID 126 146 BY SIMILARITY.  
SQ SEQUENCE 148 AA; 16265 MW; EE740FE37CFB9F0E CRC64;

Query Match 38.6%; Score 98; DB 1; Length 148;  
Best Local Similarity 41.9%; Pred. No. 6.3e-06;  
Matches 18; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKESDKYHFRIFRVLKVQRQVTDHLEHYNEMQWTTCKR 43

DB 64 NKASNDAYHSRAIOVVARAKQIVAGVNYFLDVEGMRTTCKRSQ 106

## RESULT 10

CYT\_C\_SALSC STANDARD; PRT; 146 AA.

AC O19093;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin C precursor.  
GN CST3.  
OS Saimiri sciureus (Common squirrel monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.  
NCBI\_TaxID=9521;  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE=97054523; PubMed=8898820;  
RA Wei L.H., Walker L.C., Levy E.,  
RT "Cystatin C. Icelandic-like mutation in an animal model of  
RT stroke 27:2080-2085(1996)."  
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
CC thought to serve an important physiological role as a local  
CC regulator of this enzyme activity.  
CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U52028; AAB64051.1; -.  
DR HSSP: P01034; I966.  
DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; cystatin.1.  
DR SMART: SM00043; CY.1.  
DR PROSITE: PS00287; CYSTATIN.1.  
KW Thiol protease inhibitor; Amyloid; Signal.  
FT SIGNAL 1 26 BY SIMILARITY.  
FT CHAIN 1 146 CYSTATIN C.  
FT ACT\_SITE 37 37 REACTIVE SITE.  
FT SITE 81 85 SECONDARY AREA OF CONTACT.  
FT DISULFID 99 109 BY SIMILARITY.  
FT DISULFID 123 143 BY SIMILARITY.  
SQ SEQUENCE 146 AA; 15946 MW; 08196353C0306AA3 CRC64;

Query Match 37.8%; Score 96; DB 1; Length 146;  
Best Local Similarity 43.9%; Pred. No. 1.1e-05;  
Matches 18; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKESDKYHFRIFRVLKVQRQVTDHLEHYNEMQWTTCKR 41  
DB 61 NKASNDAYHSRAIOVVARAKQIVAGVNYFLDVEGMRTTCKR 101

## RESULT 11

CYT\_CHICK STANDARD; PRT; 139 AA.

AC P01038;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin precursor (Egg-white cystatin).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Gallinae;  
OC Gallus.

OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A. MEDLINE=9008873; PubMed=2793849;  
 RA Colella R., Saksaguchi Y., Nagase H., Bird J.W.C.;  
 RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,  
 RT and tissue distribution.";  
 RL J. Biol. Chem. 264:17164-17169(1989).  
 RN [2]  
 RP SEQUENCE OF 24-139. MEDLINE=84178305; PubMed=6712597;  
 RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;  
 RT "Cystatin. Amino acid sequence and possible secondary structure.";  
 RL Biochem. J. 217:813-817(1984).  
 RN [3]  
 RP SEQUENCE OF 24-139. MEDLINE=84110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,  
 RA Machleidt W.;  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 RT of cystatin from chicken egg white.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).  
 RN [4]  
 RP CHARACTERIZATION OF PROTEIN. MEDLINE=83256421; PubMed=6409085;  
 RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,  
 RA Senter D.C., Barrett A.J.;  
 RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved  
 RT purification from egg white, characterization, and detection in  
 RT chicken serum.";  
 RL Biochem. J. 211:129-138(1983).  
 RN [5]  
 RP DISULFIDE BONDS. Grubb A., Loeiberg H., Barrett A.J.;  
 RA "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
 RT cystatin.";  
 RL FEBS Lett. 170:370-374(1984).  
 RN [6]  
 RP PHOSPHORYLATION. MEDLINE=89252033; PubMed=2721673;  
 RA Leber W., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,  
 RA Bode W.;  
 RT "The cysteine proteinase inhibitor chicken cystatin is a  
 RT phosphoprotein.";  
 RL FEBS Lett. 248:162-168(1989).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS). MEDLINE=89052676; PubMed=3191914;  
 RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshikov A.,  
 RA Brzin J., Kos J., Turk V.;  
 RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and  
 RT its possible mode of interaction with cysteine proteinases.";  
 RL EMBO J. 7:2593-2599(1988).  
 RN [8]  
 RP STRUCTURE BY NMR. MEDLINE=94087719; PubMed=8263912;  
 RA Dieckmann T., Mitschang L., Hofmann M., Kos J., Turk V.,  
 RA Auerwald E.A., Jeanlike R., Oeschkinat H.;  
 RT "The structures of native phosphorylated chicken cystatin and of a  
 RT recombinant unphosphorylated variant in solution.";  
 RL J. Mol. Biol. 234:1048-1059(1993).  
 CC -1- FUNCTION: This protein binds tightly to and inhibits a variety of  
 CC thiol proteases including ficin, papain, and cathepsins B, C, H,  
 CC and L. Although isolated from egg white, it is also present in  
 CC serum.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 CC EMBL: J05077; AAA48744.1; -.  
 DR PIR: A34456; UDCH.  
 DR PDB: 1CEW; 31-JAN-94.  
 DR PDB: 1A67; 27-MAY-98.  
 DR PDB: 1A90; 17-JUN-98.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; Cystatin; 1.  
 DR SMART: SM00043; CY; 1.  
 DR PROSITE: PS00287; CYSTATIN; 1.  
 KW Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 139  
 FT ACT\_SITE 32 32 CYSSTATIN.  
 FT SITE 76 80 REACTIVE SITE.  
 FT DISULFID 94 104 SECONDARY AREA OF CONTACT.  
 FT DISULFID 118 138  
 FT MOD\_RES 103 103  
 FT STRAND 35 36 PHOSPHORYLATION (PARTIAL).  
 FT TURN 39 40  
 FT TURN 42 51  
 FT HELIX 52 52  
 FT TURN 53 56  
 FT HELIX 57 58  
 FT TURN 81 95  
 FT STRAND 96 97  
 FT TURN 99 100  
 FT HELIX 101 108  
 FT STRAND 115 125  
 FT TURN 126 129  
 FT STRAND 130 139  
 SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;  
 Query Match 37.0%; Score 94; DB 1; Length 139;  
 Best Local Similarity 43.9%; Pred. No. 2e-05;  
 Matches 18; Conservative 9; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 NKESDDKYHFRIFRYLVKQROVTDHLEHYLVNEMQWTTCK 41  
 Db 56 NRASNDKYSRVVIVSAKQLVSGIKYILQVEIGRTTCPK 96  
 RESULT 12  
 ID CYTC HUMAN STANDARD; PRT; 146 AA.  
 AC P01034;  
 DT 21-JUN-1986 (Rel. 01, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)  
 DE (Post-gamma-globulin).  
 GN CST3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. MEDLINE=87219149; PubMed=3495457;  
 RC TISSUE=Placenta;  
 RX MEDLINE=87219149; PubMed=3495457;  
 RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;  
 RT "Molecular cloning and sequence analysis of cDNA coding for the  
 RT precursor of the human cysteine proteinase inhibitor cystatin C.";  
 RL FEBS Lett. 216:229-233(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A. MEDLINE=90303202; PubMed=2363674;  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=90303202; PubMed=2363674;  
 RA Abrahamson M., Olafsson I., Paledoclr A., Ulfvaeck M., Lundwall A.,  
 RA Jensen O., Grubb A.;  
 RT "Structure and expression of the human cystatin C gene.";

RA Biochem. J. 268:287-294(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A. (HCHWA VARIANT).  
 RC TISSUE=Brain;  
 RX MEDLINE=89235594; PubMed=2541223;  
 RA Levy E., Lopez-Olin C., Ghiso J., Gellner D., Frangione B.;  
 RT "stroke in Icelandic patients with hereditary amyloid angiopathy is  
 RT related to a mutation in the cystatin C gene, an inhibitor of  
 RT cysteine proteases.";  
 RL J. Exp. Med. 169:1771-1778(1989).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89350949; PubMed=2764935;  
 RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,  
 RA Iemura S., Sanada K.;  
 RT "The human cystatin C gene (CST3) is a member of the cystatin gene  
 RT family which is localized on chromosome 20.";  
 RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Dickinson D.P., Hewett-Emmett D., Thiesse M.;  
 RT "Acquisition of complex patterns of differential expression in  
 RT epithelial cell populations during the evolution of type 2 cystatin  
 RT genes.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stuvier G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor P.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights K., Laird G.K., Lawlor S.,  
 RA Leharvestho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurtry A.A.,  
 RA Moline S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=23388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Datchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Ueas T.B., Tomihata S., Carrino F., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko T., Boulford G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Kizylinski M.I., Skalek U., Smalins D.E.,  
 RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [8]  
 RP SEQUENCE OF 27-146.  
 RX MEDLINE=82222268; PubMed=6283552;  
 RA Grubb A., Loeffberg H.;  
 RT "Human gamma-trace, a basic microprotein: amino acid sequence and  
 RT presence in the adenohypophysis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027(1982).  
 RN [9]  
 RP SEQUENCE OF 27-73.  
 RX MEDLINE=84110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longner M., Ritonja A., Eropkin M., Borchart U.,  
 RA Machleidt W.;  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 RT of cystatin from chicken egg white.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).  
 RN [10]  
 RP SEQUENCE OF 27-76.  
 RX MEDLINE=84128015; PubMed=6365094;  
 RA Brzin J., Popovic T., Turk V.;  
 RT "Human cystatin, a new protein inhibitor of cysteine proteinases.";  
 RL Biochem. Biophys. Res. Commun. 118:103-109(1984).  
 RN [11]  
 RP DISULFIDE BONDS.  
 RA Grubb A., Loeffberg H., Barrett A.J.;  
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
 RT cystatin.";  
 RL FEBS Lett. 170:370-374(1984).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.  
 RX MEDLINE=21173909; PubMed=11276250;  
 RA Janowski R., Kozak M., Jankowska E., Grzonka Z., Grubb A.,  
 RA Abrahamson M., Jaskolski M.;  
 RT "Human cystatin C, an amyloidogenic protein, dimerizes through  
 RT three-dimensional domain swapping.";  
 RL Nat. Struct. Biol. 8:316-320(2001).  
 RN [13]  
 RP VARIANT GLN-94.  
 RX MEDLINE=92316504; PubMed=1352269;  
 RA Abrahamson M., Jonasson S., Olafsson I., Jansson O., Grubb A.;  
 RT "Hereditary cystatin C amyloid angiopathy: identification of the  
 RT disease-causing mutation and specific diagnosis by polymerase chain  
 RT reaction-based analysis.";  
 RL Hum. Genet. 89:377-380(1992).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 CC thought to serve an important physiological role as a local  
 CC regulator of this enzyme activity.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,  
 CC vas deferens, brain, thymus, and ovary and the lowest in the  
 CC submandibular gland.  
 CC -1- DISEASE: Defects in CST3 are a cause of hereditary cerebral  
 CC hemorrhage with amyloidosis (HCHWA) [MIM:105150]; also known as  
 CC cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis  
 CC Icelandic type. HCHWA is characterized by a thickening of the  
 CC cerebral arteries walls with deposition of material with the  
 CC characteristic of amyloid.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC EMBL, X05607; CAA29096.1; -

DR EMBL: X52255; CAA36497.1; -  
DR EMBL: M27891; AAA52164.1; -  
DR EMBL: M27889; AAA52164.1; JOINED.  
DR EMBL: M27890; AAA52164.1; JOINED.  
DR EMBL: X61681; CAA43856.2; -  
DR EMBL: X61682; CAA43856.2; JOINED.  
DR EMBL: X61683; CAA43856.2; JOINED.  
DR EMBL: A919564; AAK1570.1; -  
DR EMBL: AL121894; CAC05424.1; -  
DR EMBL: BC013083; AAI13083.1; -  
DR PIR: S10216; UDHU.  
DR PDB: 1G96; 06-APR-01.  
DR Genew: HGNC:2475; CST3.  
DR MIM: 604312; -  
DR MIM: 105150; -  
DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; cystatin; 1.  
DR SMART: SM00043; CY; 1.  
DR PROSITE: PS00287; CYSTATIN; 1.  
DR PROSITE: PS00287; CYSTATIN; 1.  
KW Thiol protease inhibitor; Amyloid; Signal; Disease mutation;  
KW Polymorphism; 3D-structure.  
FT SIGNAL 1 26  
FT CHAIN 27 146  
FT ACT SITE 37 37 REACTIVE SITE.  
FT SITE 81 85 SECONDARY AREA OF CONTACT.  
FT DISULFID 99 109  
FT DISULFID 123 143  
Query Match 37.0%; Score 94; DB 1; Length 146;  
Best Local Similarity 39.0%; Pred. No. 2.1e-05;  
Matches 17; Conservative 11; Mismatches 15; Indels 0; Gaps 0;  
Qy 1 NKESDDKYHPRIFRVLKVRQVTDHLEHLNVEQMWTTCOKP 43  
Db 61 NKASNDMYHSRALQVVARQIVAGVYPLDVELGRITTCRTQ 103  
RESULT 13  
CYTC\_MACNU STANDARD; PRT; 146 AA.  
AC 015092;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin C precursor.  
DE Cystatin C precursor.  
GN CST3.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97054523; Pubmed=8998820;  
RA Wei L.H., Walker L.C., Levy E.;  
RT "Cystatin C, Icelandic-like mutation in an animal model of  
cerebrovascular beta-amyloidosis.";  
RL Stroke 27:2080-2085(1996).  
-1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
thought to serve an important physiological role as a local  
regulator of this enzyme activity.  
-1- SIMILARITY: Belongs to the cystatin family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: U51912; AAB64050.1; -  
DR HSP: P01034; 1G96.

DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; cystatin; 1.  
DR SMART: SM00043; CY; 1.  
DR PROSITE: PS00287; CYSTATIN; 1.  
KW Thiol protease inhibitor; Amyloid; Signal;  
KW Polymorphism; 3D-structure.  
FT SIGNAL 1 26  
FT CHAIN 27 146  
FT ACT SITE 37 37 REACTIVE SITE.  
FT SITE 81 85 SECONDARY AREA OF CONTACT.  
FT DISULFID 99 109  
FT DISULFID 123 143  
SQ SEQUENCE 146 AA; 15857 MW; F0B3BB774A29DF26 CRC64;  
Query Match 37.0%; Score 94; DB 1; Length 146;  
Best Local Similarity 39.0%; Pred. No. 2.1e-05;  
Matches 17; Conservative 11; Mismatches 15; Indels 0; Gaps 0;  
Qy 1 NKESDDKYHPRIFRVLKVRQVTDHLEHLNVEQMWTTCOKP 43  
Db 61 NKASNDMYHSRALQVVARQIVAGVYPLDVELGRITTCRTQ 103  
RESULT 14  
CYT\_COTUA STANDARD; PRT; 116 AA.  
AC P01061;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin (Bgg-white cystatin).  
OS Coturnix coturnix japonica (Japanese quail).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Coturnix.  
OX NCBI\_TaxID=93934;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=97420480; Pubmed=9276465;  
RA Gerhartz B., Engh R.A., Mentle R., Eckerkorn C., Torguato R.,  
RA Witman J., Kolb H.J., Machleidt W., Fritz H., Auerwald E.A.;  
RT "Quail cystatin: isolation and characterization of a new member of  
RT the cystatin family and its hypothetical interaction with cathepsin  
RT B.";  
RL FBS Lett. 412:551-558(1997).  
-1- FUNCTION: This protein binds tightly to and inhibits papain and  
-1- SIMILARITY: Belongs to the cystatin family.  
-1- SIMILARITY: Belongs to the cystatin family.  
CC cathepsin B.  
CC HSP: P01038; ICEW.  
DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; cystatin; 1.  
DR PROSITE: PS00287; CYSTATIN; 1.  
KW Thiol protease inhibitor; Phosphorylation.  
FT ACT SITE 9 9 REACTIVE SITE.  
FT SITE 53 57  
FT DISULFID 71 81  
FT DISULFID 95 115  
FT MOD RES 80 80  
SQ SEQUENCE 116 AA; 13093 MW; 4824862105A2AF70 CRC64;  
Query Match 35.4%; Score 90; DB 1; Length 116;  
Best Local Similarity 39.0%; Pred. No. 5.5e-05;  
Matches 16; Conservative 11; Mismatches 14; Indels 0; Gaps 0;  
Qy 1 NKESDDKYHPRIFRVLKVRQVTDHLEHLNVEQMWTTCOK 41  
Db 33 NKASNDKYSSRVYRITSAKQIVSGIKYIMEVIGRTTCCK 73  
RESULT 15  
CSTL\_HUMAN STANDARD; PRT; 165 AA.

AC 09H114;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-like 1 precursor.  
 GN CSTL1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehaeslath M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromens A.C., Vautin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:665-871(2001).  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL: AL096677; CAC03440.2; -  
 DR HSSP; P01038; 1A90.  
 DR Genew; HGNC:15958; CSTL1.  
 DR Interpro; IPR000010; Cystatin.  
 DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
 KM Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 165 CYSTATIN-LIKE 1.  
 FT SITE 93 97 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 111 121 BY SIMILARITY.  
 FT DISULFID 134 154 BY SIMILARITY.  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 165 AA; 19312 MW; 9D6D685875DAEEA CRC64;

Query Match 34.3%; Score 87; DB 1; Length 165;  
 Best Local Similarity 27.3%; Pred. No. 0.0002;

Matches 18; Conservative 15; Mismatches 13; Indels 20; Gaps 1;  
 QY 1 NKESDDKXHPRIPRVLKVQ-----RQVTDHLEHNLVNEQWTTCC 40  
 DB 53 NNASNDIYLRYVQRLIRSQWQERVSHMLGVHINSTTDSRQLTGVGYIVTVAIGWTKCK 112  
 QY 41 KPEITN 46  
 DB 113 RNDTSN 118  
 Search completed: March 18, 2004, 14:16:11  
 Job time : 5.47619 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 27.5635 Seconds

(without alignments)  
526.560 Million cell updates/sec

Title: US-09-941-314-10

Perfect score: 254  
Sequence: 1 NKESDDKXHFRIPLVLRKVR.....EYHLNVEMQWTTCKPPTTN 46

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacterioplasmid:\*  
17: sp\_archaeoplasmid:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	100.0	138	4	Q8WXU6
2	151	59.4	139	11	Q8K5A3
3	113	44.5	141	11	Q9DAP1
4	113	44.5	141	11	Q8OZNS
5	110	43.3	103	4	Q8WXU5
6	99	39.0	140	11	Q9BFX9
7	92	36.2	112	13	Q9B8R4
8	92	36.2	112	13	Q9B8R3
9	81	31.9	109	5	Q9TY65
10	81	31.9	148	5	Q9NH95
11	80	31.5	161	5	Q16159
12	77	30.3	140	11	Q8OY72
13	76	29.9	425	3	Q12700
14	72	28.3	125	5	Q25620
15	72	28.3	146	11	Q8K397
16	72	28.3	149	11	Q9D1B1

17	71	28.0	149	11	Q8VHC1	Q8vnc1 ratu
18	71	28.0	498	5	Q16454	Q16454 caenorhabd
19	68	26.8	157	5	Q17108	Q17108 acanthopod
20	67	26.4	148	11	Q9JH84	Q9jnh84 mus muscu
21	65	25.6	167	4	Q724U8	Q724U8 homo sapien
22	64	25.2	144	13	Q8JFUS	Q8jfu5 brachydanio
23	64	25.2	284	1	Q88Z07	Q88z07 lacobacill
24	63	24.8	438	3	Q9URL8	Q9url8 candida alb
25	62	24.4	302	4	Q9H7A0	Q9h7a0 homo sapien
26	62	24.4	421	4	Q9NKS0	Q9nks0 homo sapien
27	62	24.4	693	11	Q8K145	Q8k145 mus muscu
28	62	24.4	995	4	Q9HC18	Q9hc18 homo sapien
29	61.5	24.2	587	9	Q85911	Q85911 staphylococ
30	61.5	24.2	587	9	Q85910	Q85910 staphylococ
31	61	24.0	127	5	P90698	P90698 brugia mala
32	61	24.0	787	11	Q8B182	Q8b182 mus muscu
33	60	23.6	204	4	Q8TCY7	Q8tcy7 homo sapien
34	60	23.6	275	4	Q8WY62	Q8wy62 homo sapien
35	60	23.6	357	4	Q72465	Q72465 homo sapien
36	59	23.2	429	3	Q12628	Q12628 kiuyveromyc
37	58.5	23.0	133	5	Q8WYB6	Q8wyb6 ixodes scap
38	58.5	23.0	246	16	Q980Y6	Q980y6 mycoplasma
39	58.5	23.0	544	6	Q28548	Q28548 ovis aries
40	58.5	23.0	891	5	Q81B61	Q81b61 plasmodium
41	58	22.8	159	4	Q8TD53	Q8td53 homo sapien
42	58	22.8	167	11	Q9OWL5	Q9owl5 mus muscu
43	58	22.8	540	5	Q8T9G8	Q8t9g8 drosophila
44	58	22.8	610	4	Q9UJ02	Q9uj02 homo sapien
45	58	22.8	631	4	Q9NXR3	Q9nxr3 homo sapien

# ALIGNMENTS

RESULT 1	
Q8WXU6	PRELIMINARY; PRT; 138 AA.
AC Q8WXU6	
DT 01-MAR-2002 (TREMREL. 20, Created)	
DT 01-MAR-2002 (TREMREL. 20, Last sequence update)	
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)	
DE SC13	
OS Homo sapiens (Human)	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
OX NCBI_Taxid=9606;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA Hamil K.G., Liu Q., Zhang Y.-L., French P.S., Hall S.H.;	
RT "SC13: A novel epididymal specific member of the cystatin family."	
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AF35480; AAL7191.1; -	
DR GO; GO:0004869; P:cysteine protease inhibitor activity; IEA.	
DR InterPro; IPR000010; Cystatin.	
DR Pfam; PF00031; Cystatin; 1.	
DR SMART; SM00043; CY; 1.	
DR SEQUENCE 138 AA; 16506 MW; E49440ACA3585C64 CRC64;	
QY	
Query Match	100.0%; Score 254; DB 4; Length 138;
Best Local Similarity	100.0%; Pred. No. 1.8e-26;
Matches	46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	
1 NKESDDKXHFRIPLVLRKVR.....EYHLNVEMQWTTCKPPTTN 46	
56 NKESDDKXHFRIPLVLRKVR.....EYHLNVEMQWTTCKPPTTN 101	
RESULT 2	
Q8K5A3	PRELIMINARY; PRT; 139 AA.
AC Q8K5A3	
DT 01-OCT-2002 (TREMREL. 22, Created)	
DT 01-OCT-2002 (TREMREL. 22, Last sequence update)	

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE Cystatin 11.  
 GN CST11.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RA Hamil K.G., Hall S.H.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF501290; AAM21709.1;  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 SQ SEQUENCE 139 AA; 16686 MW; E1E36DB786B4D08C CRC64;  
 QY Query Match 59.4%; Score 151; DB 11; Length 139;  
 Best Local Similarity 58.1%; Pred. No. 1.5e-12;  
 Matches 25; Conservative 12; Mismatches 6; Indels 0; Gaps 0;  
 DB 1 NKESDDKYHFRIFRVLVKVRQVTDHLEVHNVEMQWTTCKR 43  
 56 NKSEDLNFRVILRIKIKEMQTNHMEHITVEMQRTCLKTE 98  
 RESULT 3  
 Q9DAP1 PRELIMINARY; PRT; 141 AA.  
 AC Q9DAP1;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE 1700006C19R1K protein.  
 GN 1700006C19R1K.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; Tissue=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 Arai K., Hara A., Fukunishi Y., Kono H., Maehi J., Fukuda S.,  
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 Schiraldi L.M., Staudt F., Suzuki K., Tomita M., Wagner L., Washio T.,  
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 Blake J., Botfield D., Boujarda N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberey P.,  
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 Wyszynski A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 Yashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 RL EMBL; AK005665; BAB24175.1;  
 DR HSSP; P01038; ICEW.  
 DR MGD; MGI:191654; 1700006C19R1K.  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 SQ SEQUENCE 141 AA; 16811 MW; C20FA0DB8B1A378C CRC64;

Query Match 44.5%; Score 113; DB 11; Length 141;  
 Best Local Similarity 51.2%; Pred. No. 2.1e-07;  
 Matches 21; Conservative 8; Mismatches 12; Indels 0; Gaps 0;  
 DB 1 NKESDDKYHFRIFRVLVKVRQVTDHLEVHNVEMQWTTCKR 41  
 56 NKASNDLNFVRVVDILKSQEQITDSLEYEVNIAFTMCKR 96  
 RESULT 4  
 Q80ZNS PRELIMINARY; PRT; 141 AA.  
 AC Q80ZNS;  
 DT 01-JUN-2003 (TReMBLrel. 24, Created)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE RIKEN cDNA 1700006C19 gene.  
 DE Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Testicle;  
 RA Strauberg R.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC048681; AAH48681.1;  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR InterPro; IPR003243; Cystatin\_C/M.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR ProDom; PD001231; Cystatin\_C/M; 1.  
 DR SMART; SM00043; CY; 1.  
 SQ SEQUENCE 141 AA; 16825 MW; C20FA0DB8A84951F CRC64;  
 QY Query Match 44.5%; Score 113; DB 11; Length 141;  
 Best Local Similarity 51.2%; Pred. No. 2.1e-07;  
 Matches 21; Conservative 8; Mismatches 12; Indels 0; Gaps 0;  
 DB 1 NKESDDKYHFRIFRVLVKVRQVTDHLEVHNVEMQWTTCKR 41  
 56 NKASNDLNFVRVVDILKSQEQITDSLEYEVNIAFTMCKR 96  
 RESULT 5  
 Q8WXUS PRELIMINARY; PRT; 103 AA.  
 AC Q8WXUS;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE SC13delta.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;  
 RT "SC13: A novel epidiymal specific member of the cystatin family."  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF35481; AAL71992.1;  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;  
 QY Query Match 43.3%; Score 110; DB 4; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-07;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 1 NKESDDKYHFRIFRVLVKVRQ 21

Db 56 NKSSDDKYHFRIFVLRVQVTDHLEHNLVEMQWTTCCOK 76

## RESULT 6

Q9E9X9 PRELIMINARY; PRT: 140 AA.

AC Q9E9X9 01-JUN-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Cystatin C.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxId=10090;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c;  
 RX MEDLINE=21010502; PubMed=11144350;

RA Taupin P.-J., Ray J., Fischer W.H., Suh S.T., Hakansson K., Grubb A.,  
 RA Gage F.H.;

RT "GF-2-Responsive neural stem cell proliferation requires CCG, a novel  
 RT autocrine/paracrine cofactor.";

RL Neuron 28:385-397(2000).  
 DR EMBL; AF311741; AAC40283.1; -

DR HSSP; P01034; 1G96.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; Cystatin; 1.

DR SMART; SM00043; CY; 1.

DR PROSITE; PS00287; CYSTATIN; 1

FT CHAIN 21 140 CYSTATIN C.  
 FT VARIANT 16 16 A -> G.  
 FT VARIANT 84 84 L -> P.

SO SEQUENCE 140 AA; 15517 MW; 3A563406D58D785 CRC64;

Query Match 39.0%; Score 99; DB 11; Length 140;  
 Best Local Similarity 43.2%; Pred. No. 1.6e-05;  
 Matches 19; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

## RESULT 7

Q96SR4 PRELIMINARY; PRT: 112 AA.

AC Q96SR4 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Cystatin (Fragment).  
 OS Acipenser sinensis (Chinese sturgeon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;  
 OC Acipenser.

NCBI\_TaxId=61970;

RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Bai J., Lao H., Ye X., Li Y., Lou J.;

RT "Molecular cloning and sequence analysis of cystatin cDNA from two  
 RT species of sturgeons.";

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF334610; AAK16731.1; -

DR HSSP; P01038; 1A90.  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.

FT NON TER 1 1  
 SO SEQUENCE 112 AA; 12231 MW; 48CEBFD8A08C00 CRC64;

Query Match 36.2%; Score 92; DB 13; Length 112;  
 Best Local Similarity 43.9%; Pred. No. 0.00011;  
 Matches 18; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Q9 1 NKSSDDKYHFRIFVLRVQVTDHLEHNLVEMQWTTCCOK 41

Db 28 NKASNDWYIHRVSKVQVQVAVGIVTVQWGRSCKR 68

## RESULT 8

Q96SR3 PRELIMINARY; PRT: 112 AA.

AC Q96SR3 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Cystatin (Fragment).  
 OS Acipenser schrenckii (Amur sturgeon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;  
 OC Acipenser.

NCBI\_TaxId=111304;

RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Bai J., Lao H., Ye X., Li Y., Lou J.;

RT "Molecular cloning and sequence analysis of cystatin cDNA from two  
 RT species of sturgeons.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF334611; AAK16732.1; -

DR HSSP; P01038; 1A90.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; Cystatin; 1.

DR PROSITE; PS00287; CYSTATIN; 1.

FT NON TER 1 1  
 SO SEQUENCE 112 AA; 12231 MW; 48CEBFD8A08C00 CRC64;

Query Match 36.2%; Score 92; DB 13; Length 112;  
 Best Local Similarity 43.9%; Pred. No. 0.00011;  
 Matches 18; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Q9 1 NKSSDDKYHFRIFVLRVQVTDHLEHNLVEMQWTTCCOK 41

Db 28 NKASNDWYIHRVSKVQVQVAVGIVTVQWGRSCKR 68

## RESULT 9

Q9TY65 PRELIMINARY; PRT: 109 AA.

AC Q9TY65 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Antigen maltose binding protein (Fragment).  
 OS Onchocerca volvulus.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Onchocerca.  
 NCBI\_TaxId=6282;

RN [1]  
 RP SEQUENCE FROM N.A.

RC MEDLINE=94336252; PubMed=8058358;

RA Bradley J.B.;

RT "Antigenicity of IgG antibody responses to cloned Onchocerca volvulus  
 RT antigens in microfiladermia positive individuals from Bameraldas  
 RT Province, Ecuador.";

RL Parasite Immunol. 16:201-209(1994).  
 DR EMBL; S71364; AAC60509.1; -

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

```

DR InterPro; IPR000010; Cystatin.
DR Pfam; PF000031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT NON_TER 1
SQ SEQUENCE 109 AA; 12701 MW; 366B5EF14D8082B CRC64;

Query Match
Best Local Similarity 34.1%; Score 81; DB 5; Length 109;
Matches 14; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

QY 1 NKESDDKHFRIFRVLYKQROYTDHLEHLNVEHQWTTCK 41
DB 24 NQSNDEYHLMPIKLVKSSQVAGVYKXMDVQVARSCK 64

RESULT 10
Q9NH95 PRELIMINARY; PRT; 148 AA.
ID Q9NH95
AC Q9NH95;
DT 01-OCT-2000 (TrEMBLrel. 15; Created)
DT 01-OCT-2000 (TrEMBLrel. 15; Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24; Last annotation update)
DE Lb-cystatin.
OS Litomosoides sigmodontis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Litomosoides.
OX NCBI_TaxID=42156;
RN [1]
RP SEQUENCE FROM N.A.
RA Pfaff A.W., Hoffmann W.H., Taylor D.W., Schulz-Key H.;
RT "Characterization and immunological properties of a cysteine protease
RT inhibitor of the filarial parasite Litomosoides sigmodontis.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229173; AAF35886.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT CHAIN 25 148
SQ SEQUENCE 148 AA; 16686 MW; 2350AA89CA5339C9 CRC64;

Query Match
Best Local Similarity 31.9%; Score 81; DB 5; Length 148;
Matches 14; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKESDDKHFRIFRVLYKQROYTDHLEHLNVEHQWTTCK 41
DB 62 NQSNDAVHLMPIKLVKSSQVAGVYKXMDVQVARSCK 102

RESULT 11
O16159 PRELIMINARY; PRT; 161 AA.
ID O16159
AC O16159;
DT 01-JAN-1998 (TrEMBLrel. 05; Created)
DT 01-JAN-1998 (TrEMBLrel. 05; Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24; Last annotation update)
DE Cystatin-type cysteine proteinase inhibitor.
DE BM-CPI-2.
OS Brugia malayi (Filarial nematode worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Brugia.
OX NCBI_TaxID=6279;
RN [1]
RP SEQUENCE FROM N.A.
RA Gregory W.F., Blaxter M.L., Maizels R.M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Gregory W.F., Maizels R.M.;
RL "Two distinct cystatin-type cysteine protease inhibitors from the

```

```

RT parasitic nematode Brugia malayi.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF015263; AAB69857.1; -
DR EMBL; AF177193; AAD51086.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
SQ SEQUENCE 161 AA; 18406 MW; 8081351EBE226EB5 CRC64;

Query Match
Best Local Similarity 31.5%; Score 80; DB 5; Length 161;
Matches 14; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKESDDKHFRIFRVLYKQROYTDHLEHLNVEHQWTTCK 41
DB 73 NQSNDEYHLMPIKLVKSSQVAGVYKXMDVQVARSCK 113

RESULT 12
O80Y72 PRELIMINARY; PRT; 140 AA.
ID O80Y72
AC O80Y72;
DT 01-JUN-2003 (TrEMBLrel. 24; Created)
DT 01-JUN-2003 (TrEMBLrel. 24; Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25; Last annotation update)
DE Cystatin-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=22386257; PubMed=12477932;
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marisla K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Straube R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048646; A448646.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 140 AA; 16199 MW; 32633E98C4697DA0 CRC64;

Query Match
Best Local Similarity 30.3%; Score 77; DB 11; Length 140;
Matches 15; Conservative 14; Mismatches 15; Indels 0; Gaps 0;

```

QY 1 NKSDDKYHFRIFRVLVKVRQVTDHLEHYHNVEMQWTTCCOK 44  
 DB 56 NNASNDTYLVQVQKLIQSQMQLTGVEXLVTVKIGRTCKKNET 99

## RESULT 13

ID 012700 PRELIMINARY; PRT; 425 AA.

AC 012700;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Exo-1,3-beta-glucanase/1,3-beta-D-glucan glucanohydrolase  
 DE (EC 3.2.1.58).  
 OS Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.  
 OX NCBI\_TaxID=27300;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 26077;  
 RX MEDLINE=99154256; PubMed=10029988;  
 RA Escoban P., del Rey F., Vazquez De Aldana C.R.;  
 RT "Cloning and characterization of 1,3-beta-glucanase-encoding genes  
 from non-conventional yeasts."  
 RL Yeast 15:91-109(1999).  
 DR EMBL; Z46871; CA86951.1; -.  
 DR HSSP; P29717; ICZ1.  
 DR GO: GO:0004338; F:Glucan 1,3-beta-glucosidase activity; IEA.  
 DR GO: GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . , IEA.  
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro: IPR001547; Glyco\_hydro\_5.  
 DR PROSITE; PS00659; GLYCOSYL\_HYDROL\_F5; 1.  
 KM Glycosidase; Hydrolase.  
 SQ SEQUENCE 425 AA; 49127 MW; 57F063ABE2FBF274 CRC64;

Query Match 29.9%; Score 76; DB 3; Length 425;

Best Local Similarity 28.8%; Pred. No. 0.066; Mismatches 11; Indels 16; Gaps 2;

QY 6 DKYHFRIFRVLVKVRQVTDHLEHYHNVEMQWTTCCOK 41  
 DB 277 DHHYVQFSAGELQSIDHTIVACMGWDAKKEHYHNVAGESALTDCAK 328

## RESULT 14

ID 025620 PRELIMINARY; PRT; 125 AA.

AC 025620;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Beta-galactosidase.  
 DE Beta-galactosidase.  
 GN LACZ.  
 OS Onchocerca volvulus.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Onchocerca.  
 OX NCBI\_TaxID=6282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92042729; PubMed=1840605;  
 RA Chandaashekar R., Masood K., Alvarez R.M., Ogunrinade A.F., Lujan R.,  
 RA Richards F.O.Jr., Wei G.J.;  
 RT "Molecular cloning and characterization of recombinant parasite  
 antigens for immunodiagnosis of onchocerciasis."  
 RL J. Clin. Invest. 88:1460-1466(1991).  
 DR EMBL; M60279; AAA29418.1; -.  
 DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.

SO SEQUENCE 125 AA; 14369 MW; C73D5D9C7495A3CB CRC64;

Query Match 28.3%; Score 72; DB 5; Length 125;

Best Local Similarity 31.7%; Pred. No. 0.065; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKSDDKYHFRIFRVLVKVRQVTDHLEHYHNVEMQWTTCCOK 41  
 DB 40 NEQSDENHLMPIKLVKSSQVAVGAVKIMDVQVANSQCCK 80

## RESULT 15

ID 08K397 PRELIMINARY; PRT; 146 AA.

AC 08K397;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE RIKEN cDNA 110017E11 gene (fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Strauberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC027680; AA827680.1; -.  
 DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 FT NON TER 1  
 SQ SEQUENCE 146 AA; 16380 MW; 9D77BB9A063A5C4 CRC64;

Query Match 28.3%; Score 72; DB 11; Length 146;

Best Local Similarity 31.7%; Pred. No. 0.076; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKSDDKYHFRIFRVLVKVRQVTDHLEHYHNVEMQWTTCCOK 41  
 DB 57 NMGSDSLYFRDTRKVIDAKYQLVAGIKYVLTLDIESIECKK 97

Search completed: March 18, 2004, 14:21:28  
 Job time : 27.5635 secs





CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (1) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (1) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX

SO Sequence 24 AA;

Query Match 100.0%; Score 138; DB 5; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-14;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTTCKPKE 24  
 Db 1 RQVTDHLEHYHNVEMQWTTCKPKE 24

RESULT 2

AAU79860  
 ID AAU79860 standard; peptide; 46 AA.

AC AAU79860;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #8.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

OS WO200220567-A2.

PN 14-MAR-2002.

PD 29-AUG-2001; 2001WO-US026868.

PF 01-SEP-2000; 2000US-0230230P.

PR (ZYMO) ZYMOGENETICS INC.

PA Holloway JL, Gao Z, Bishop PD;

PI WPI; 2002-383044/41.

DR Novel isolated mammalian cystatin-8 polypeptide useful for promoting

PT spermatogenesis, and inhibiting cancer procoagulant protein which leads

PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 97; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis; modulating seminal  
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (1) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (1) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX

SO Sequence 46 AA;

Query Match 100.0%; Score 138; DB 5; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-14;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTTCKPKE 24  
 Db 20 RQVTDHLEHYHNVEMQWTTCKPKE 43

RESULT 3

AAU79863  
 ID AAU79863 standard; peptide; 49 AA.

AC AAU79863;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #11.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

OS WO200220567-A2.

PN 14-MAR-2002.

PD 29-AUG-2001; 2001WO-US026868.

PF 01-SEP-2000; 2000US-0230230P.

PR (ZYMO) ZYMOGENETICS INC.

PA Holloway JL, Gao Z, Bishop PD;

PI WPI; 2002-383044/41.

DR Novel isolated mammalian cystatin-8 polypeptide useful for promoting

PT spermatogenesis, and inhibiting cancer procoagulant protein which leads

PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 97-98; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis; modulating seminal  
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (1) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (1) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX  
 SQ Sequence 49 AA;

Query Match 100.0%; Score 138; DB 5; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 7,7e-14;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTCCKPE 24  
 |||  
 DB 23 RQVTDHLEHYHNVEMQWTCCKPE 46

RESULT 4  
 AAU79864 standard; peptide; 52 AA.  
 XX  
 AC AAU79864;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #12.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 98; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis; modulating seminal  
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX Sequence 52 AA;

Query Match 100.0%; Score 138; DB 5; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 8,3e-14;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTCCKPE 24  
 |||  
 DB 19 RQVTDHLEHYHNVEMQWTCCKPE 42

RESULT 5  
 AAU79866 standard; peptide; 59 AA.  
 XX  
 ID AAU79866;

AC AAU79866;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #14.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 99; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis; modulating seminal  
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX Sequence 59 AA;

Query Match 100.0%; Score 138; DB 5; Length 59;  
 Best Local Similarity 100.0%; Pred. No. 9,6e-14;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTTCKPPE 24  
ID 1 RQVTDHLEHYHNVEMQWTTCKPPE 24

RESULT 6  
AAU79865 standard; peptide; 80 AA.

AAU79865;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #13.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

MO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
spermatogenesis, and inhibiting cancer procoagulant protein which leads  
to inhibition of thrombotic events associated with cancer.

Claim 2; Page 98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)  
polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
protein in an individual and thus inhibiting the thrombotic events  
associated with cancer; promoting spermatogenesis, modulating seminal  
fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
motility and fertilisation; and as antigenic peptides to generate  
antibodies. Zcys8 is useful as research reagent for characterising sites  
of interaction between Zcys8 and its receptor. Zcys8 is useful in  
enhancing fertilisation during assisted reproduction in humans and in  
animals. Anti-(I) antibodies are useful to screen biological samples like  
blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
presence of Zcys8. The antibodies are also useful to isolate large  
quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
The polynucleotide encoding (I) is useful to detect and to localise the  
expression of a Zcys8 gene in a biological sample and Zcys8  
oligonucleotide probes are useful for in vivo diagnosis. The  
polynucleotide encoding (I) is useful in determining whether a subject's  
chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
copy number changes, insertions, deletions, restriction site changes and  
rearrangements and genetic alterations that inactivate the Zcys8 gene.  
This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

Sequence 80 AA;

Query Match 100.0%; Score 138; DB 5; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.4e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTTCKPPE 24  
ID 1 RQVTDHLEHYHNVEMQWTTCKPPE 24

DB 22 RQVTDHLEHYHNVEMQWTTCKPPE 45

RESULT 7  
AAU79853 standard; protein; 115 AA.

AAU79853;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #1.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
sperm motility; fertilisation; antigenic fragment.

Homo sapiens.

MO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
spermatogenesis, and inhibiting cancer procoagulant protein which leads  
to inhibition of thrombotic events associated with cancer.

Claim 2; Page 94; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)  
polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
protein in an individual and thus inhibiting the thrombotic events  
associated with cancer; promoting spermatogenesis, modulating seminal  
fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
motility and fertilisation; and as antigenic peptides to generate  
antibodies. Zcys8 is useful as research reagent for characterising sites  
of interaction between Zcys8 and its receptor. Zcys8 is useful in  
enhancing fertilisation during assisted reproduction in humans and in  
animals. Anti-(I) antibodies are useful to screen biological samples like  
blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
presence of Zcys8. The antibodies are also useful to isolate large  
quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
The polynucleotide encoding (I) is useful to detect and to localise the  
expression of a Zcys8 gene in a biological sample and Zcys8  
oligonucleotide probes are useful for in vivo diagnosis. The  
polynucleotide encoding (I) is useful in determining whether a subject's  
chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
copy number changes, insertions, deletions, restriction site changes and  
rearrangements and genetic alterations that inactivate the Zcys8 gene.  
This sequence represents an antigenic fragment of human cystatin-8  
(Zcys8)

Sequence 115 AA;

Query Match 100.0%; Score 138; DB 5; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.1e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTTCKPPE 24  
ID 1 RQVTDHLEHYHNVEMQWTTCKPPE 24

DB 52 RQVTDHLEHYHNVEMQWTTCKPPE 75

RESULT 8  
AAU79854

ID AAU79854 standard; protein; 117 AA.  
 XX  
 AC AAU79854;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #2.  
 XX  
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic fragment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 94-95; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 XX  
 SQ Sequence 117 AA;  
 XX  
 QY Query Match 100.0%; Score 138; DB 5; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-13;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 RQVTDHLEHYLNVEMQWTTCKRPE 24  
 Db 54 RQVTDHLEHYLNVEMQWTTCKRPE 77  
 XX  
 RESULT 9  
 AAU79852  
 ID AAU79852 standard; protein; 137 AA.  
 XX  
 AC AAU79852;  
 XX

XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8).  
 XX  
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 DR N-PSDB; ABK49522.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 93-94; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 137 AA;  
 XX  
 QY Query Match 100.0%; Score 138; DB 5; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-13;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 RQVTDHLEHYLNVEMQWTTCKRPE 24  
 Db 74 RQVTDHLEHYLNVEMQWTTCKRPE 97  
 XX  
 RESULT 10  
 AAU79862  
 ID AAU79862 standard; peptide; 33 AA.  
 XX  
 AC AAU79862;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX

DE Human cystatin-8 (Zcys8) antigenic fragment #10.  
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.  
 XX Homo sapiens.  
 OS  
 PN MO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001MO-US026668.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 97; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 33 AA;  
 Query Match 96.4%; Score 133; DB 5; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 2,9e-13;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 QVTDHLEVHLNVEQMWTTCQKE 24  
 Db 1 QVTDHLEVHLNVEQMWTTCQKE 23  
 RESULT 11  
 AAU79867  
 ID AAU79867 standard; peptide; 48 AA.  
 XX  
 AC AAU79867;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #15.  
 XX  
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KM sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001MO-US026668.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 99; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 48 AA;  
 Query Match 55.8%; Score 77; DB 5; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 0.00021;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 12 NVEMQWTTTCQKE 24  
 Db 1 NVEMQWTTTCQKE 13  
 RESULT 12  
 ADD46708  
 ID ADD46708 standard; protein; 142 AA.  
 XX  
 AC ADD46708;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Rat Protein AAC36317, SEQ ID NO 12393.  
 XX  
 KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KM chronic constriction injury; CCI, spared nerve injury; SNI; Chung.  
 OS Rattus norvegicus.  
 XX  
 PN MO2003016475-A2.

XX 27-FEB-2003.  
 PD 14-AUG-2002; 2002MO-US025765.  
 XX 14-AUG-2001; 2001US-0312147P.  
 XX 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 PI WPI; 2003-268312/26.  
 DR GENBANK; AAC36317.  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 PS Claim 1; Page; 1017pp; English.  
 XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 142 AA;  
 Query Match 51.4%; Score 71; DB 7; Length 142;  
 Best Local Similarity 45.5%; Pred. No. 0.0064;  
 Matches 10; Conservative 8; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 QVTDHLEHYHNVEMQWTCCKP 23  
 Db 77 QITDRMEYHIDVOISNSCKRP 98  
 RESULT 13  
 ADD46704 standard; protein; 142 AA.  
 AC ADD46704;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Rat Protein AAC36317, SEQ ID NO 12389.  
 XX  
 KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.  
 OS  
 XX WO2003016475-A2.  
 PN  
 XX 27-FEB-2003.  
 PD  
 XX 14-AUG-2002; 2002MO-US025765.  
 PF  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 PI WPI; 2003-268312/26.  
 DR GENBANK; AAC36317.  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 PS Claim 1; Page; 1017pp; English.  
 XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 142 AA;  
 Query Match 51.4%; Score 71; DB 7; Length 142;  
 Best Local Similarity 45.5%; Pred. No. 0.0064;  
 Matches 10; Conservative 8; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 QVTDHLEHYHNVEMQWTCCKP 23  
 Db 77 QITDRMEYHIDVOISNSCKRP 98  
 RESULT 14  
 AAE02404 standard; protein; 142 AA.  
 AC AAE02404;  
 XX  
 DT 10-AUG-2001 (first entry)  
 XX

DE Murine cystatin-related epididymal specific protein (CRS).

XX Murine; Cystatin T; zcy33; cystatin-related epididymal specific protein;  
 KM CRS; inhibitor; cysteine proteinase; male reproductive tissue; testis;  
 KM spermatogenesis; therapy; reproductive disorder.

XX Mus musculus.  
 OS  
 PN US6235708-B1.  
 XX  
 PD 22-MAY-2001.

PF 01-NOV-1999; 99US-00431480.  
 XX  
 PR 20-NOV-1998; 98US-0109217P.  
 PR 28-SEP-1999; 99US-0156382P.

PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Feldhaus AL;  
 XX  
 DR WPI; 2001-342846/36.

XX Cystatin T polypeptides are useful for modulating spermatogenesis and  
 PT studying, diagnosing and treating reproductive disorders.

XX  
 PS Disclosure; Col 45-46; 32pp; English.

XX The present invention relates to cystatin T (also known as zcy33) DNA and  
 CC protein sequences. Cystatin T is testis specific and is homologous to  
 CC cystatin-related epididymal specific gene (CRS) and type 2 cystatins.  
 CC Cystatins inhibit cystine proteinases and are found with male  
 CC reproductive tissues and secretions. Cystatin T sequence is useful for  
 CC modulating spermatogenesis and studying, diagnosing and treating  
 CC reproductive disorders. The present sequence is murine cystatin-related  
 CC epididymal specific (CRS) protein

XX  
 SQ Sequence 142 AA;

QY Query Match 45.7%; Score 63; DB 4; Length 142;  
 Db Best Local Similarity 40.9%; Pred. No. 0.11; Mismatches 5; Indels 0; Gaps 0;  
 Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
 2 QVTDHLEHYHNVEMQWTTCKP 23  
 77 QITDRMEYQIDVQISRSNCKKP 98

RESULT 15  
 AAE04433  
 ID AAE04433 standard; protein; 142 AA.  
 XX  
 AC AAE04433;  
 XX  
 DT 04-SEP-2001 (first entry)  
 XX  
 DE Mouse cystatin-related epididymal specific (CRS) protein.  
 XX  
 KW Mouse; cystatin T; zcy33; testis specific; spermatogenesis modulator;  
 KW cystatin-related epididymal specific gene; CRS; gene therapy;  
 KW sperm production; antifertility.  
 XX  
 OS Mus musculus.  
 XX  
 PN US6245529-B1.  
 XX  
 PD 12-JUN-2001.  
 XX  
 PF 17-JUL-2000; 2000US-00617302.  
 XX  
 PR 20-NOV-1998; 98US-0109217P.  
 PR 28-SEP-1999; 99US-0156382P.  
 PR 01-NOV-1999; 99US-00431480.

XX (ZYMO ) ZYMOGENETICS INC.  
 PA  
 XX Holloway JL, Feldhaus AL;  
 PI  
 XX WPI; 2001-407271/43.  
 DR  
 XX  
 XX New polynucleotides encoding testis-specific cystatin-like protein  
 PT cystatin T, useful in gene therapy for modulating cystatin T activity,  
 PT particularly for modulating spermatogenesis, or enhancing sperm  
 PT production or fertility.

XX  
 PS Disclosure; Col 47-48; 33pp; English.

XX The present sequence is mouse cystatin-related epididymal specific (CRS)  
 CC protein which is homologous to mouse testis specific cystatin T (also  
 CC known as zcy33). The cystatin T polynucleotide is useful in gene therapy  
 CC applications, where it is desired to increase or inhibit cystatin T  
 CC activity. It is also useful for producing cystatin T polypeptide, as well  
 CC as for detecting the expression of a cystatin T gene in a biological  
 CC sample. The cystatin T is useful for modulating spermatogenesis, and may  
 CC be used to study or modulate that function in vitro or in vivo  
 CC systems. In particular, it is also useful for enhancing sperm production,  
 CC increasing the number of viable sperm in a sample, or enhancing  
 CC fertilisation

XX  
 SQ Sequence 142 AA;

QY Query Match 45.7%; Score 63; DB 4; Length 142;  
 Db Best Local Similarity 40.9%; Pred. No. 0.11; Mismatches 5; Indels 0; Gaps 0;  
 Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
 2 QVTDHLEHYHNVEMQWTTCKP 23  
 77 QITDRMEYQIDVQISRSNCKKP 98

Search completed: March 18, 2004, 14:15:00  
 Job time : 22.7619 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 6.14286 Seconds  
(without alignments)  
201.701 Million cell updates/sec

Title: US-09-941-314-11  
Perfect score: 138  
Sequence: 1 RQVTDHLEHYLNVEMQWTTCKPK 24

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues  
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patente AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/6C\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	45.7	142	3	US-09-431-480-3
2	63	45.7	142	3	US-09-617-302-3
3	60	43.5	141	3	US-09-431-480-2
4	60	43.5	141	3	US-09-617-302-2
5	56	40.6	142	3	US-09-431-480-4
6	56	40.6	142	3	US-09-617-302-4
7	52	37.7	112	4	US-08-849-303-19
8	52	37.7	112	4	US-08-849-303-16
9	52	37.7	118	4	US-09-775-932-24
10	50	36.2	68	4	US-08-858-207A-492
11	50	36.2	714	2	US-08-472-534-3
12	49	35.5	140	4	US-09-886-319A-46
13	49	35.5	140	4	US-09-886-319A-48
14	49	35.5	145	2	US-08-832-535-2
15	49	35.5	145	3	US-09-019-485-2
16	49	35.5	145	3	US-09-019-485-3
17	49	35.5	145	3	US-09-431-480-9
18	49	35.5	145	3	US-09-617-302-9
19	49	35.5	145	4	US-09-528-436B-2
20	49	35.5	178	2	US-08-791-522-1
21	49	35.5	178	3	US-09-314-777-1
22	48.5	35.1	3546	4	US-09-679-279-13
23	48	34.8	111	4	US-08-849-303-26
24	48	34.8	116	4	US-09-775-932-16
25	48	34.8	120	4	US-09-775-932-2
26	48	34.8	120	6	5432264-4
27	48	34.8	139	2	US-08-791-522-4

28	48	34.8	139	3	US-09-314-777-4	Sequence 4, Appli
29	48	34.8	139	4	US-08-849-303-15	Sequence 15, Appl
30	48	34.8	145	2	US-08-832-535-11	Sequence 11, Appl
31	48	34.8	146	2	US-08-791-522-3	Sequence 3, Appli
32	48	34.8	146	3	US-08-744-138-3	Sequence 4, Appli
33	48	34.8	146	3	US-09-019-485-4	Sequence 3, Appli
34	48	34.8	146	3	US-09-314-777-3	Sequence 3, Appli
35	48	34.8	146	3	US-09-431-480-6	Sequence 6, Appli
36	48	34.8	146	3	US-09-617-302-6	Sequence 6, Appli
37	48	34.8	146	4	US-09-241-376-3	Sequence 3, Appli
38	48	34.8	146	4	US-09-528-436B-3	Sequence 3, Appli
39	48	34.8	146	4	US-09-886-319A-47	Sequence 47, Appl
40	48	34.8	146	4	US-09-940-497-3	Sequence 3, Appli
41	48	34.8	146	4	US-09-976-594-17	Sequence 37, Appl
42	48	34.8	146	4	US-08-849-303-17	Sequence 17, Appl
43	48	34.8	146	5	PCT-US95-07135-9	Sequence 9, Appli
44	48	34.8	146	6	5432264-6	Patent No. 5432264
45	48	34.8	443	4	US-09-107-532A-7119	Sequence 7119, Ap

## ALIGNMENTS

```

RESULT 1
US-09-431-480-3
; Sequence 3, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; PILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; EARLIER FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-431-480-3

Query Match          45.7%; Score 63; DB 3; Length 142;
Best Local Similarity 40.9%; Pred. No. 0.0092;
Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY      2 QVTDHLEHYLNVEMQWTTCKPK 23
      |||:||||:|:|:|:|:|:|
Db      77 QITRMEYQIDVOIRSNCKPK 98

RESULT 2
US-09-617-302-3
; Sequence 3, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; EARLIER FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22

```

SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-617-302-3

Query Match 45.7%; Score 63; DB 3; Length 142;  
Best Local Similarity 40.9%; Pred. No. 0.0092;  
Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEQMWTTCOKP 23  
DB 77 QITDMEYQIDVQISRSNCKRP 98

RESULT 3  
US-09-431-480-2  
Sequence 2, Application US/09431480  
Patent No. 6235708  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72  
CURRENT APPLICATION NUMBER: US/09/431,480  
CURRENT FILING DATE: 1999-11-01  
EARLIER APPLICATION NUMBER: 60/109,217  
EARLIER FILING DATE: 1998-11-20  
EARLIER APPLICATION NUMBER: 60/156,382  
EARLIER FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 141  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-431-480-2

Query Match 43.5%; Score 60; DB 3; Length 141;  
Best Local Similarity 52.4%; Pred. No. 0.028;  
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEQMWTTCOKP 22  
DB 76 QITDSELYLEVNIARTWCKK 96

RESULT 4  
US-09-617-302-2  
Sequence 2, Application US/09617302  
Patent No. 6245529  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
APPLICANT: Feldhaus, Andrew  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72 C1  
CURRENT APPLICATION NUMBER: US/09/617,302  
CURRENT FILING DATE: 2000-07-17  
PRIOR APPLICATION NUMBER: 09/431,480  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: 60/109,217  
PRIOR FILING DATE: 1998-11-20  
PRIOR APPLICATION NUMBER: 60/156,382  
PRIOR FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 141  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-617-302-2

Query Match 43.5%; Score 60; DB 3; Length 141;  
Best Local Similarity 52.4%; Pred. No. 0.028;  
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEQMWTTCOKP 22  
DB 76 QITDSELYLEVNIARTWCKK 96

RESULT 5  
US-09-431-480-4  
Sequence 4, Application US/09431480  
Patent No. 6235708  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72  
CURRENT APPLICATION NUMBER: US/09/431,480  
CURRENT FILING DATE: 1999-11-01  
EARLIER APPLICATION NUMBER: 60/109,217  
EARLIER FILING DATE: 1998-11-20  
EARLIER APPLICATION NUMBER: 60/156,382  
EARLIER FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-431-480-4

Query Match 40.6%; Score 56; DB 3; Length 142;  
Best Local Similarity 50.0%; Pred. No. 0.13;  
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEQMWTTCOKP 23  
DB 77 QVTNLELYLIDVEIARSDCRKP 98

RESULT 6  
US-09-617-302-4  
Sequence 4, Application US/09617302  
Patent No. 6245529  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
APPLICANT: Feldhaus, Andrew  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72 C1  
CURRENT APPLICATION NUMBER: US/09/617,302  
CURRENT FILING DATE: 2000-07-17  
PRIOR APPLICATION NUMBER: 09/431,480  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: 60/109,217  
PRIOR FILING DATE: 1998-11-20  
PRIOR APPLICATION NUMBER: 60/156,382  
PRIOR FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-617-302-4

Query Match 40.6%; Score 56; DB 3; Length 142;  
Best Local Similarity 50.0%; Pred. No. 0.13;  
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEQMWTTCOKP 23  
DB 77 QVTNLELYLIDVEIARSDCRKP 98

```

RESULT 7
US-08-849-303-19
Sequence 19, Application US/08849303
Patent No. 6680424
GENERAL INFORMATION:
APPLICANT: Atkinson, Howard J.
APPLICANT: McPherson, Michael J.
APPLICANT: Urwin, Peter E.
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,303
FILING DATE: 21-MAY-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1321-1-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-849-303-19

Query Match      38.4% Score 53; DB 4; Length 127;
Best Local Similarity 41.7%; Pred.No. 0.35;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY      1 RGVTDHEHYHNVNQWTTCKKPE 24
       : : : : : : : : : : : :
Db      61 KQLVAGINYVLDEVEMRTTCTKSQ 84

RESULT 8
US-08-849-303-16
Sequence 16, Application US/08849303
Patent No. 6680424
GENERAL INFORMATION:
APPLICANT: Atkinson, Howard J.
APPLICANT: McPherson, Michael J.
APPLICANT: Urwin, Peter E.
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:

```

```

      / MEDIUM TYPE: Floppy disk
      / COMPUTER : IBM PC compatible
      / OPERATING SYSTEM: PC-DOS/MS-DOS
      / SOFTWARE : Patentin Release #1.0, Version #1.30
      / CURRENT APPLICATION DATA:
      / APPLICATION NUMBER: US/08/849,303
      / FILING DATE: 21-MAY-1997
      / CLASSIFICATION: 514
      / ATTORNEY/AGENT INFORMATION:
      / NAME: Jackson Beq., David A.
      / REGISTRATION NUMBER: 26,742
      / REFERENCE/DOCKET NUMBER: 1321-1-003
      / TELECOMMUNICATION INFORMATION:
      / TELEPHONE: 201-487-5800
      / TELEFAX: 201-343-1684
      / TELEX: 133521
      / INFORMATION FOR SEQ ID NO: 16:
      / SEQUENCE CHARACTERISTICS:
      / LENGTH: 112 amino acids
      / TYPE: amino acid
      / STRANDEDNESS: single
      / TOPOLOGY: linear
      / MOLECULE TYPE: protein
      / HYPOTHEICAL: NO
      /
US-08-849-303-16

Query Match          37.7%; Score 52; DB 4; Length 112;
Best Local Similarity 41.7%; Pred.No. 0.44;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      1 RQVTDHLEYNLVEMQWTTCKKPS 24
DB      47 KQVSGMNYFLDVELGRITCTKSQ 70

RESULT 9
US-09-775-932-24
Sequence 24, Application US/09775932
Patent No. 6534477
GENERAL INFORMATION:
APPLICANT: University of British Columbia
TITLE OF INVENTION: Production and use of Modified Cytatins
FILE REFERENCE: 58069
CURRENT APPLICATION NUMBER: US/09/775,932
CURRENT FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: CA99/00717
PRIOR FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: 60/095,503
PRIOR FILING DATE: 1998-08-05
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 24
LENGTH: 118
TYPE: PRT
ORGANISM: Bos taurus
US-09-775-932-24

Query Match          37.7%; Score 52; DB 4; Length 118;
Best Local Similarity 41.7%; Pred. No. 0.47;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      1 RQVTDHLEYNLVEMQWTTCKKPS 24
DB      53 KQVSGMNYFLDVELGRITCTKSQ 76

RESULT 10
US-08-858-207A-492
Sequence 492, Application US/08858207A
Patent No. 6348328
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John

```

APPLICANT: Knowles, David  
APPLICANT: Nicholas, Richard  
APPLICANT: Stodola, Robert  
TITLE OF INVENTION: No. 634832861 Compounds  
NUMBER OF SEQUENCES: 552  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/858,207A  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/017670  
FILING DATE: 14-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50475  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 492:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 63483286  
US-08-858-207A-492

Query Match 36.2%; Score 50; DB 4; Length 68;  
Best Local Similarity 37.9%; Pred. No. 0.51;  
Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Qy 1 RQVTDH-----LEYHNVEMQWTTCKP 23  
Db 30 RQWTDHPNGDFMETPLNTQFDWNGIRKP 58

RESULT 11  
US-08-472-534-3  
Sequence 3, Application US/08472534  
Patent No. 5919620  
GENERAL INFORMATION:  
APPLICANT: Hamel, Josee  
APPLICANT: Brodeur, Bernard R  
APPLICANT: Martin, Denis  
TITLE OF INVENTION: HEAT SHOCK PROTEIN HSP72 FROM  
NUMBER OF SEQUENCES: 6  
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,534  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr, James F  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B1ovac-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9030  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 714 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-472-534-3

Query Match 36.2%; Score 50; DB 2; Length 714;  
Best Local Similarity 37.9%; Pred. No. 9.1;  
Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Qy 1 RQVTDH-----LEYHNVEMQWTTCKP 23  
Db 301 RQWTDHPNGDFMETPLNTQFDWNGIRKP 329

RESULT 12  
US-09-886-319A-46  
Sequence 46, Application US/09886319A  
Patent No. 6586185  
GENERAL INFORMATION:  
APPLICANT: Wolf, Eckard  
APPLICANT: Werner, Sabine  
APPLICANT: Halle, Jorn-Peter  
APPLICANT: Regenbogen, Johannes  
APPLICANT: Goppelt, Andreas  
TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound  
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically  
FILE REFERENCE: 50125/014002  
CURRENT APPLICATION NUMBER: US/09/886,319A  
PRIOR FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: US 60/222,081  
PRIOR FILING DATE: 2000-08-01  
PRIOR APPLICATION NUMBER: DE 10030149.5  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 46  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-886-319A-46

Query Match 35.5%; Score 49; DB 4; Length 140;  
Best Local Similarity 41.7%; Pred. No. 1.8;  
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 RQVTDHLEYHNVEMQWTTCKPE 24  
Db 74 KQLVAGVYFLDVEWGRITCTKSG 97

RESULT 13  
US-09-886-319A-48  
Sequence 48, Application US/09886319A  
Patent No. 6586185  
GENERAL INFORMATION:  
APPLICANT: Wolf, Eckard

APPLICANT: Werner, Sabine  
APPLICANT: Halle, Jorn-Peter  
APPLICANT: Regendogen, Johannes  
APPLICANT: Goppelt, Andreas  
TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound  
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically  
FILE REFERENCE: 50125/014002  
CURRENT APPLICATION NUMBER: US/09/886,319A  
CURRENT FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: US 60/222,081  
PRIOR FILING DATE: 2000-08-01  
PRIOR APPLICATION NUMBER: DE 10030149.5  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 48  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-886-319A-48

Query Match 35.5%; Score 49; DB 4; Length 140;  
Best Local Similarity 41.7%; Pred. No. 1.8;  
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 1 QVTDHLEHYLNVEMQWTTCKPE 24  
Db 74 KQLVAGVNYPLDVEIMGRTCTCKSQ 97

RESULT 14  
US-08-832-535-2  
Sequence 2, Application US/08832535  
Patent No. 5919658  
GENERAL INFORMATION:  
APPLICANT: NI, JIAN  
APPLICANT: LI, HAODONG  
APPLICANT: YU, GUO-LIANG  
APPLICANT: GENTZ, REINER L  
TITLE OF INVENTION: HUMAN CYSTATIN F  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
STREET: 9410 KEY WEST AVENUE  
CITY: ROCKVILLE  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/832,535  
FILING DATE: 03-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KIMBALL, PAUL C.  
REGISTRATION NUMBER: 34,610  
REFERENCE/DOCKET NUMBER: PP265  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 994-1744  
TELEFAX: (201) 994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 145 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-832-535-2

Query Match 35.5%; Score 49; DB 2; Length 145;  
Best Local Similarity 47.6%; Pred. No. 1.9;  
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 QVTDHLEHYLNVEMQWTTCK 22  
Db 81 QIVKGLKYMVEIGRTTCK 101

RESULT 15  
US-09-019-485-2  
Sequence 2, Application US/09019485  
Patent No. 6066617  
GENERAL INFORMATION:  
APPLICANT: LI, Haodong  
APPLICANT: YU, Guo-Liang  
APPLICANT: GENTZ, Reiner  
APPLICANT: NI, Jian  
TITLE OF INVENTION: Cystatin F  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/019,485  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Robert H.  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PP265P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 3013098439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 145 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-019-485-2

Query Match 35.5%; Score 49; DB 3; Length 145;  
Best Local Similarity 47.6%; Pred. No. 1.9;  
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 QVTDHLEHYLNVEMQWTTCK 22  
Db 81 QIVKGLKYMVEIGRTTCK 101

Search completed: March 18, 2004, 14:25:36  
Job time: 6.14286 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 14.5238 Seconds

(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-11

Perfect score: 138  
Sequence: 1 RQVTDHLEHYLNEMQWTTCKPE 24Scoring table: BLOSUM62  
Gapox 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubppaa/US09C\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubppaa/US10C\_NEW\_PUB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	138 100.0	24	US-09-941-314-11	Sequence 11, Appl
2	138 100.0	46	US-09-941-314-10	Sequence 10, Appl
3	138 100.0	49	US-09-941-314-13	Sequence 13, Appl
4	138 100.0	52	US-09-941-314-14	Sequence 14, Appl
5	138 100.0	59	US-09-941-314-15	Sequence 15, Appl
6	138 100.0	80	US-09-941-314-16	Sequence 16, Appl
7	138 100.0	115	US-09-941-314-3	Sequence 3, Appl
8	138 100.0	117	US-09-941-314-4	Sequence 4, Appl
9	138 100.0	137	US-09-941-314-2	Sequence 2, Appl
10	133 96.4	33	US-09-941-314-12	Sequence 12, Appl
11	77 55.8	48	US-09-941-314-17	Sequence 17, Appl
12	60 43.5	203	US-09-893-737-164	Sequence 164, Appl
13	56 40.6	138	US-10-239-663-37	Sequence 37, Appl
14	55 39.9	37	US-09-864-761-40966	Sequence 40966, A
15	55 39.9	145	US-09-740-638-2	Sequence 2, Appl

16	55 39.9	145 13	US-10-006-467-2	Sequence 2, Appl
17	55 39.9	145 14	US-10-235-148-2	Sequence 2, Appl
18	55 39.9	165 9	US-09-740-638-5	Sequence 5, Appl
19	55 39.9	165 13	US-10-006-467-5	Sequence 5, Appl
20	55 39.9	165 14	US-10-235-148-5	Sequence 14, Appl
21	53 38.4	127 8	US-08-849-303-19	Sequence 19, Appl
22	52 37.7	112 8	US-08-849-303-16	Sequence 16, Appl
23	52 37.7	118 9	US-09-775-932-24	Sequence 24, Appl
24	51.5 37.5	678 12	US-10-282-122A-61129	Sequence 61129, A
25	50 36.2	145 14	US-10-168-425-14	Sequence 14, Appl
26	49 35.5	140 14	US-10-376-564-46	Sequence 46, Appl
27	49 35.5	140 14	US-10-376-564-48	Sequence 48, Appl
28	49 35.5	145 14	US-10-329-428-2	Sequence 2, Appl
29	49 35.5	167 10	US-09-746-783-197	Sequence 197, Appl
30	49 35.5	178 9	US-09-969-834-1	Sequence 1, Appl
31	48 34.8	111 8	US-08-849-303-26	Sequence 26, Appl
32	48 34.8	116 9	US-09-775-932-16	Sequence 16, Appl
33	48 34.8	120 9	US-09-775-932-2	Sequence 2, Appl
34	48 34.8	139 8	US-08-849-303-15	Sequence 15, Appl
35	48 34.8	139 9	US-09-969-834-4	Sequence 4, Appl
36	48 34.8	146 8	US-08-849-303-17	Sequence 17, Appl
37	48 34.8	146 9	US-09-940-497-3	Sequence 3, Appl
38	48 34.8	146 9	US-09-969-834-3	Sequence 3, Appl
39	48 34.8	146 14	US-10-329-428-3	Sequence 3, Appl
40	48 34.8	146 14	US-10-376-564-47	Sequence 47, Appl
41	48 34.8	438 12	US-10-282-122A-57752	Sequence 57752, A
42	47 34.1	273 14	US-10-156-761-12874	Sequence 12874, A
43	46 33.3	122 9	US-09-775-932-10	Sequence 20, Appl
44	46 33.3	142 8	US-08-849-303-20	Sequence 20, Appl
45	46 33.3	142 9	US-09-940-497-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-09-941-314-11  
Sequence 11, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
PRIOR FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-11

Query Match 100.0%; Score 138; DB 9; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.6e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYLNEMQWTTCKPE 24  
DB 1 RQVTDHLEHYLNEMQWTTCKPE 24

RESULT 2  
US-09-941-314-10  
Sequence 10, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein

FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 46  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-10

Query Match  
Best Local Similarity 100.0%; Score 138; DB 9; Length 46;  
Pred. No. 3,2e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKP 24  
Db 20 RQVTDHLEHYHLNVEMQWTTCKP 43

RESULT 3  
US-09-941-314-13  
Sequence 13, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-13

Query Match  
Best Local Similarity 100.0%; Score 138; DB 9; Length 49;  
Pred. No. 3,4e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKP 24  
Db 23 RQVTDHLEHYHLNVEMQWTTCKP 46

RESULT 4  
US-09-941-314-14  
Sequence 14, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 52  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-14

Query Match  
Best Local Similarity 100.0%; Score 138; DB 9; Length 52;  
Pred. No. 3,7e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKP 24  
Db 19 RQVTDHLEHYHLNVEMQWTTCKP 42

RESULT 5  
US-09-941-314-15  
Sequence 15, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16  
LENGTH: 59  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-16

Query Match  
Best Local Similarity 100.0%; Score 138; DB 9; Length 59;  
Pred. No. 4,2e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKP 24  
Db 1 RQVTDHLEHYHLNVEMQWTTCKP 24

RESULT 6  
US-09-941-314-15  
Sequence 15, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-15

Query Match  
Best Local Similarity 100.0%; Score 138; DB 9; Length 80;  
Pred. No. 5,7e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKP 24  
Db 22 RQVTDHLEHYHLNVEMQWTTCKP 45

RESULT 7  
US-09-941-314-3  
Sequence 3, Application US/09941314  
Patent No. US20020142396A1

```
/ GENERAL INFORMATION:
/ APPLICANT: ZymoGenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941.314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 115
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-3

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 115;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEWMQWTTCKP 24
Db 52 RQVTDHLEHYHLNVEWMQWTTCKP 75

RESULT 8
US-09-941-314-4
/ Sequence 4, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: ZymoGenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941.314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 117
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-4

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 117;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEWMQWTTCKP 24
Db 54 RQVTDHLEHYHLNVEWMQWTTCKP 77
```

```
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-2

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 137;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEWMQWTTCKP 24
Db 74 RQVTDHLEHYHLNVEWMQWTTCKP 97

RESULT 10
US-09-941-314-12
/ Sequence 12, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: ZymoGenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941.314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12
/ LENGTH: 133
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-12

Query Match
Best Local Similarity 100.0%; Score 133; DB 9; Length 133;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEWMQWTTCKP 24
Db 1 QVTDHLEHYHLNVEWMQWTTCKP 23

RESULT 11
US-09-941-314-17
/ Sequence 17, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: ZymoGenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941.314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17
/ LENGTH: 148
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-17

Query Match
Best Local Similarity 100.0%; Score 77; DB 9; Length 148;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NVEMQWTTCKP 24
Db 1 NVEMQWTTCKP 13
```

```
RESULT 12
US-09-893-737-164
; Sequence 164, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 164
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-164

Query Match      43.5%; Score 60; DB 9; Length 203;
Best Local Similarity 52.4%; Pred. No. 0.61;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY      2 QVTDHLEHYHNVEMQWTTCK 22
      |||||:::|::|::|:|
Db      115 QVTDSELYEYIEVKIARTICK 115

RESULT 13
US-10-239-663-37
; Sequence 37, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GPS0018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-663-37

Query Match      40.6%; Score 56; DB 14; Length 138;
Best Local Similarity 47.6%; Pred. No. 1.6;
Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40986
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121894.14
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EST HUMAN HIT: AM665128.1, EVALUATE 1.10e-02
; OTHER INFORMATION: SWISSPROT HIT: O88969, EVALUATE 1.00e-04
US-09-864-761-40986

Query Match      39.9%; Score 55; DB 9; Length 37;
Best Local Similarity 50.0%; Pred. No. 0.57;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
```

```

RESULT 15
US-09-740-638-2
: Sequence 2, Application US/09740638
: Patent No. US20020006656A1
: GENERAL INFORMATION:
: APPLICANT: Holloway, James L.
: TITLE OF INVENTION: Zcy45: A Member of the Cystatatin
: FILE REFERENCE: Superfamily
: CURRENT APPLICATION NUMBER: US/09/740,638
: CURRENT FILING DATE: 2000-12-18
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 145
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-740-638-2

```

Query Match	39.9%	Score 55;	DB 9;	Length 145;
Best Local Similarity	40.9%	Pred. No. 2.4;		
Matches	9;	Conservative	7;	Mismatches 6;
				Indels 0;
				Gaps 0;

```
Qy      1 RQVTDHLEHYHLNVEMQWTTCCQK 22
          ||:| :||: || |:
Db      72 RQLTGV EYIVTVKIGWTKCKR 93
```

Search completed: March 18, 2004, 14:30:54  
Job time : 14.5238 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 4.61905 Seconds  
(without alignments)  
499.799 Million cell updates/sec

Title: US-09-941-314-11  
Perfect score: 138  
Sequence: 1 RQVTDHLEHYHLNEMQWTTQCKPE 24

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR 78:\*  
2: p1r1:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	45.7	139	2 A45361	cystatin-related e
2	53	38.4	120	2 S10587	cystatin C - rat
3	53	38.4	127	2 S07085	cystatin C precurs
4	53	38.4	1779	2 T23130	hypothetical prote
5	52	37.7	112	1 UD80	cystatin - bovine
6	50	36.2	232	2 T28369	probable uracil-DN
7	50	36.2	588	2 C95252	L-fucose isomerase
8	48	36.2	588	2 A99717	cystatin - puff ad
9	48	34.8	111	2 A28793	cystatin precursor
10	48	34.8	139	1 UDCH	cystatin C precurs
11	48	34.8	146	1 UDHU	cystatin C precurs
12	47.5	34.4	519	2 B88322	protein C47D12.8 l
13	47.5	34.4	519	2 T19988	hypothetical prote
14	47.5	34.4	1308	2 B75198	DNA helicase relat
15	47	34.1	456	2 B84960	adenylosuccinate l
16	47	34.1	555	2 S39953	isocitrate lyase (
17	46.5	33.7	371	2 B96614	hypothetical prote
18	46	33.3	133	2 JC4536	cystatin precursor
19	46	33.3	142	2 A47142	cystatin D precurs
20	46	33.3	355	2 T34405	hypothetical prote
21	46	33.3	677	1 SYECWT	histidinol-phospha
22	46	33.3	677	1 SYECWT	methionine-CRNA 1i
23	46	33.3	677	2 C85839	methionine-CRNA 1i
24	46	33.3	677	2 H90993	methionine-CRNA 1i
25	45.5	33.0	652	2 S58666	serine/threonine-p
26	45.5	33.0	1640	2 H88094	protein F399.2 (i
27	45	32.6	140	2 A36163	cystatin C precurs
28	45	32.6	454	2 T27249	hypothetical prote
29	45	32.6	455	2 T00856	pectate lyase (EC

30	45	32.6	482	2 T15333	hypothetical prote
31	45	32.6	675	2 AF0185	methionine-CRNA 1i
32	45	32.6	913	2 G64110	hypothetical prote
33	44.5	32.2	2055	2 T31617	hypothetical prote
34	44	31.9	142	2 AC1642	hypothetical prote
35	44	31.9	436	2 AC1279	hypothetical prote
36	44	31.9	447	2 KGB0L1	kininogen, LMW I P
37	44	31.9	447	2 D71812	udp-n-acetylmuram
38	44	31.9	497	2 A84956	kininogen, HMW I P
39	44	31.9	621	1 KGB0H1	protein F182.16 l
40	44	31.9	1075	2 C96682	DNA-directed RNA p
41	44	31.9	1478	2 S78131	hypothetical prote
42	43.5	31.5	419	2 T25577	protein K03H1.2 (i
43	43.5	31.5	1131	2 F88570	hypothetical prote
44	43.5	31.5	1451	2 S41025	fiber - human aden
45	43	31.2	325	2 D37476	

## ALIGNMENTS

## RESULT 1

A45361  
cystatin-related epididymal specific protein - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C:Accession: A45361  
R:Corwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.  
Mol. Endocrinol. 6, 1653-1664, 1992  
A>Title: The CREB gene: a unique testis-regulated gene related to the cystatin family is  
A:Reference number: A45361; MUID:93078799; PMID:1280328  
A:Accession: A45361  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-139 <COR>  
A:Cross-references: GB:949926; MID:g260492; PIDN:AAC35390.1; PID:g260493  
A>Note: sequence extracted from NCBI backbone (NCBI:118813)  
C:Superfamily: cystatin; cystatin homology  
F:28-139/Domain: cystatin homology <CYS>

Query Match 45.7%; Score 63; DB 2; Length 139;  
Best Local Similarity 40.9%; Pred. No. 0.029;  
Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNEMQWTTQCKP 23  
Db 74 QITDREYQIDVOISRSNCKP 95

## RESULT 2

S10587  
cystatin C - rat  
C:Species: Rattus sp. (rat)  
C>Date: 21-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 16-Jul-1999  
C:Accession: S10587  
R:Bernard, F.; Banaud, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Brillard, M.; Gauch  
Biochem. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990  
A>Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylat  
A:Reference number: S10587; MUID:90380276; PMID:2400577  
A:Accession: S10587  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-120 <ESN>  
A>Note: 43-Asn was also found  
A>Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala  
C:Superfamily: cystatin; cystatin homology  
F:9-120/Domain: cystatin homology <CYS>

Query Match 38.4%; Score 53; DB 2; Length 120;  
Best Local Similarity 41.7%; Pred. No. 0.81;  
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNEMQWTTQCKP 24



L-fucose isomerase [imported] - Streptococcus pneumoniae (strain TIGR4)  
 C/Species: Streptococcus pneumoniae  
 C/Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
 C/Accession: C95252  
 R/Retellin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
 on, J.D.; Umeyama, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
 nson, T.; Hickey, E.K.; Holt, I.E.  
 Science 293, 498-506, 2001  
 A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
 A./Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
 A/Reference number: A95000; MUID:21357209; PMID:11463916  
 A/Accession: C95252  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-588 <KUR>  
 A/Cross-references: GB:AE005672; PIDN:AAK76212.1; PID:914973669; GSPDB:GN00164; TIGR:SPA  
 A/Experimental source: strain TIGR4  
 C/Genetics:  
 A/Gene: SP2158  
 C/Superfamily: isomerase fucI

Query Match 36.2%; Score 50; DB 2; Length 588;  
 Best Local Similarity 37.9%; Pred. No. 13;  
 Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

OY 1 RQVTDH-----LEVHLNVEKQMTTCCKP 23  
 Db 301 RQWTDHPNGDFMETPLNTQPDWNGIRKP 329

RESULT 8  
 A99717  
 L-fucose isomerase (EC 5.3.1.25) [imported] - Streptococcus pneumoniae (strain R6)  
 C/Species: Streptococcus pneumoniae  
 C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 02-Nov-2001  
 C/Accession: A99717  
 R/Hoskins, J.A.; Albhorn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAlhnen, S.; M  
 y, P.; Sun, P.M.; Winkler, M.E.  
 J. Bacteriol. 183, 5709-5717, 2001  
 A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
 A./Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
 A/Reference number: A97872; MUID:21429245; PMID:11544234  
 A/Accession: A99717  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-588 <KUR>  
 A/Cross-references: GB:AE007317; PIDN:AAU00766.1; PID:915459665; GSPDB:GN00174  
 C/Genetics:  
 A/Gene: fucI  
 C/Superfamily: isomerase fucI  
 C/Keywords: intramolecular oxidoreductase; isomerase

Query Match 36.2%; Score 50; DB 2; Length 588;  
 Best Local Similarity 37.9%; Pred. No. 13;  
 Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

OY 1 RQVTDH-----LEVHLNVEKQMTTCCKP 23  
 Db 301 RQWTDHPNGDFMETPLNTQPDWNGIRKP 329

RESULT 9  
 A28793  
 Cystatin - puff adder  
 C/Species: Bittis arietans (puff adder)  
 C/Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 30-Sep-1993  
 C/Accession: A28793  
 R/Rittonja, A.; Evans, H.J.; Machleidt, W.; Barrett, A.J.  
 Biochem. J. 246, 799-802, 1987  
 A./Title: Amino acid sequence of a cystatin from venom of the African puff adder (Bittis a  
 A/Reference number: A28793; MUID:88076661; PMID:3500714  
 A/Accession: A28793

A/Molecule type: protein  
 A/Residues: 1-111 <RT>  
 C/Superfamily: cystatin; cystatin homology

Query Match 34.8%; Score 48; DB 2; Length 111;  
 Best Local Similarity 42.9%; Pred. No. 4.2;  
 Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 2 QVTDHLEVHLNVEKQMTTCCKP 22  
 Db 47 QVSVGVKXVLMMEBLKTTCKK 67

RESULT 10  
 UDCH  
 Cystatin precursor - chicken  
 N/Alternate names: cystatin 1; cysteine proteinase inhibitor; egg-white cystatin  
 C/Species: Gallus gallus (chicken)  
 C/Date: 03-Aug-1984 #sequence\_revision 12-Apr-1996 #text\_change 29-Oct-1999  
 C/Accession: A34456; A01274; S01461; S48159; S04008; JN0789  
 R/Colella, R.; Sakaguchi, Y.; Nagase, H.; Bird, J.W.C.  
 J. Biol. Chem. 264, 17164-17169, 1989  
 A./Title: Chicken egg white cystatin. Molecular cloning, nucleotide sequence, and tissue  
 A/Reference number: A34456; MUID:90008873; PMID:2793849  
 A/Accession: A34456  
 A/Molecule type: mRNA  
 A/Residues: 1-139 <COL>  
 A/Cross-references: GB:J05077; NID:9211714; PIDN:AAA48744.1; PID:9211715  
 R/Schwabe, C.; Anastasi, A.; Crow, H.; McDonald, J.K.; Barrett, A.J.  
 Biochem. J. 217, 813-817, 1984  
 A./Title: Cystatin. Amino acid sequence and possible secondary structure.  
 A/Reference number: A01274; MUID:84178305; PMID:6712557  
 A/Accession: A01274  
 A/Molecule type: protein  
 A/Residues: 24-139 <SCH>  
 R/Turk, V.; Brzin, J.; Longer, M.; Rittonja, A.; Eropkin, M.; Borchardt, U.; Machleidt, W.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983  
 A./Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatrin  
 A/Reference number: S01461; MUID:84110059; PMID:6662498  
 A/Accession: S01461  
 A/Molecule type: protein  
 A/Residues: 24-139 <TUR>  
 R/Anastasi, A.; Brown, M.A.; Kembhavi, A.A.; Nicklin, M.J.H.; Sayers, C.A.; Sunter, D.C.  
 Biochem. J. 211, 129-138, 1983  
 A./Title: Cystatin, a protein inhibitor of cysteine proteinases. Improved purification fr  
 A/Reference number: A37514; MUID:83256421; PMID:6409085  
 A/Contents: annotation; characterization of protein  
 R/Grubb, A.; Lofberg, H.; Barrett, A.J.  
 FEBS Lett. 170, 370-374, 1984  
 A./Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
 A/Reference number: S01462  
 A/Contents: annotation; disulfide bonds  
 R/Auerswald, E.A.; Naegler, D.K.; Schulze, A.J.; Engh, R.A.; Genenger, G.; Machleidt, W.;  
 Eur. J. Biochem. 224, 407-415, 1994  
 A./Title: Production, inhibitory activity, folding and conformational analysis of an N-ter  
 A/Reference number: S48159; MUID:95010016; PMID:7925354  
 A/Accession: S48159  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 24-139 <AUE>  
 R/Lieber, B.; Kriegelstein, K.; Hennechen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.  
 FEBS Lett. 248, 162-168, 1989  
 A./Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.  
 A/Reference number: S04008; MUID:89252033; PMID:2721673  
 A/Accession: S04008  
 A/Molecule type: protein  
 A/Residues: 97-114 <LAB>  
 R/Colella, R.; Bird, J.W.C.  
 Gene 130, 175-181, 1993  
 A./Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping to  
 A/Reference number: JN0789; MUID:93366172; PMID:8359684  
 A/Accession: JN0789  
 A/Molecule type: DNA

A/Residues: 1-139 <CO2>  
 A/Cross-references: GB:M5725  
 A/Note: authors failed to translate the codon for residue 115-Tyr  
 C/Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases  
 C/Genetics:  
 A/Gene: Csn  
 A/Introns: 76/3; 114/3  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein  
 F/1-23/Domain: signal sequence #status predicted <SIG>  
 F/24-139/Product: cystatin, long form #status experimental <CYLF>  
 F/30-139/Domain: cystatin homology <CYS>  
 F/33-139/Product: cystatin, short form #status experimental <CYSF>  
 F/76-80/Region: inhibitory #status predicted  
 F/94-104/Region: disulfide bonds #status experimental  
 F/103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 34.8%; Score 48; DB 1; Length 139;  
 Best Local Similarity 45.5%; Pred. No. 5.4;  
 Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

CY 1 RQVTDHEHNVEMQWTCCK 22  
 DB 75 RQVSGIKYIQVEIRGTTCPK 96

RESULT 11  
 UDHU  
 cystatin C precursor [validated] - human  
 N/Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma  
 C/Species: Homo sapiens (man)  
 C/Date: 06-Jul-1992 #sequence revision 31-Mar-1991 #text change 08-Dec-2000  
 C/Accession: S10216; S00004; J10095; A33400; S02751; A01270; A25434; S12288; A32732; A60  
 R/Abrahamson, M.; Olafsson, I.; Palsdottir, A.; Ulfvback, M.; Lundwall, A.; Jensen, O.  
 Biochem. J. 268, 287-294, 1990  
 A/Title: Structure and expression of the human cystatin C gene.  
 A/Reference number: S10216; MUID:90303202; PMID:2363674  
 A/Accession: S10216  
 A/Molecule type: DNA  
 A/Residues: 1-146 <AB1>  
 A/Cross-references: EMBL:X52255; NID:G30257; PIDN:CAA36497.1; PID:G296643  
 R/Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.  
 FEBS Lett. 216, 229-233, 1987  
 A/Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the  
 A/Reference number: S00004; MUID:87219149; PMID:3494557  
 A/Accession: S00004  
 A/Molecule type: mRNA  
 A/Residues: 1-146 <AB2>  
 A/Cross-references: EMBL:X05607; NID:G30371; PIDN:CAA29096.1; PID:G755738  
 R/Levy, E.; Lopez-Otin, C.; Ghiso, J.; Gellner, D.; Frangione, B.  
 J. Exp. Med. 169, 1771-1778, 1989  
 A/Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a  
 A/Reference number: J10095; MUID:85235594; PMID:2541223  
 A/Accession: J10095  
 A/Molecule type: DNA  
 A/Residues: 1-146 <LEU>  
 A/Cross-references: GB:X61681; NID:G30367; PIDN:CAA43856.2; PID:G4490944  
 A/Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit  
 e)  
 R/Saitoh, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada,  
 Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989  
 A/Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which  
 A/Reference number: A33400; MUID:89350949; PMID:2764935  
 A/Accession: A33400  
 A/Molecule type: DNA  
 A/Residues: 1-24, 'T', 26-146 <SAT>  
 A/Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:G181385; PIDN:AAA52164.1; PID:G  
 R/Ghiso, J.; Cowan, N.; Frangione, B.  
 Biol. Chem. Hoppe-Seyler 369, 205-208, 1988  
 A/Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron  
 A/Reference number: S02751; MUID:89076507; PMID:3264504  
 A/Accession: S02751  
 A/Molecule type: DNA

A/Residues: 82-119 <GH2>  
 A/Cross-references: EMBL:M2769  
 A/Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown i  
 R/Grubb, A.; Lofberg, H.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982  
 A/Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in the  
 A/Reference number: A01270; MUID:8222268; PMID:6283552  
 A/Accession: A01270  
 A/Molecule type: protein  
 A/Residues: 27-131, 'S', 133-146 <GRU>  
 R/Ghiso, J.; Jensen, O.; Frangione, B.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986  
 A/Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland ty  
 A/Reference number: A25434; MUID:86206076; PMID:3517880  
 A/Accession: A25434  
 A/Molecule type: protein  
 A/Residues: 37-93, 'Q', 95-146 <GHI>  
 R/Turk, V.; Brzin, J.; Longier, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983  
 A/Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatir  
 A/Reference number: S01461; MUID:84110059; PMID:6662498  
 A/Accession: S12288  
 A/Molecule type: protein  
 A/Residues: 27-73 <TUR>  
 R/Brzin, J.; Popovic, T.; Turk, V.  
 Biochem. Biophys. Res. Commun. 118, 103-109, 1984  
 A/Title: Human cystatin, a new protein inhibitor of cysteine proteinases.  
 A/Reference number: A32732; MUID:84128015; PMID:6365094  
 A/Accession: A32732  
 A/Molecule type: protein  
 A/Residues: 27-76 <BE2>  
 R/Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jensen, O.; Grubb, A.  
 Scand. J. Clin. Lab. Invest. 50, 85-93, 1990  
 A/Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst  
 A/Reference number: A60552; MUID:90193615; PMID:2315647  
 A/Accession: A60552  
 A/Molecule type: protein  
 A/Residues: 27-49, 'XX', 52-64 <OLA>  
 A/Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d  
 e defective gene is not present in CSF but is found instead in amyloid deposits  
 R/Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.  
 Biol. Chem. Hoppe-Seyler 371, 575-580, 1990  
 A/Title: Different forms of human cystatin C.  
 A/Reference number: S10607; MUID:91025625; PMID:2222856  
 A/Accession: S10607  
 A/Molecule type: protein  
 A/Residues: 27-53 <POP>  
 A/Experimental source: urine, kidney disease  
 A/Note: truncated forms with amino ends at positions 35 and 36 of the precursor were als  
 R/Grubb, A.; Lofberg, H.; Barrett, A.J.  
 FEBS Lett. 170, 370-374, 1984  
 A/Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
 A/Reference number: S01462  
 A/Contents: annotation; disulfide bonds  
 R/Berti, P.J.; Storer, A.C.  
 Biochem. J. 302, 411-416, 1994  
 A/Title: Local pH-dependent conformational changes leading to proteolytic susceptibility  
 A/Reference number: S55305; MUID:94379969; PMID:8092921  
 A/Accession: S55305  
 A/Molecule type: preliminary  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 27-49, 106-146 <BER>  
 C/Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl  
 e patients with certain autoimmune diseases.  
 C/Comment: This protein is an inhibitor of cysteine proteinases and may serve an import  
 C/Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorr  
 C/Genetics:  
 A/Gene: GST3  
 A/Cross-references: GDB:119817; OMIM:105150  
 A/Map position: 20p11.2-20p11.2  
 A/Introns: 81/3; 119/3  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 2.85714 Seconds  
(without alignments)  
437.389 Million cell updates/sec

Title: US-09-941-314-11  
Perfect score: 138

Sequence: 1 RQYTDHLEHYHLNEMQWTCCKRPE 24

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	100.0	137	CS11_HUMAN	Q9H112 homo sapien
2	91	65.9	139	CS11_MOUSE	Q94269 mus musculu
3	71	51.4	142	CS18_RAT	O88969 rattus norv
4	63	45.7	142	CS18_MOUSE	P32766 mus musculu
5	56	40.6	142	CS18_HUMAN	O60676 homo sapien
6	55	39.9	165	CS18_HUMAN	Q9H114 homo sapien
7	53	38.4	127	CS18_RAT	P14841 rattus norv
8	52	37.7	148	CS18_MOUSE	P01035 mus musculu
9	50	36.2	146	CS18_MOUSE	O19093 rattus norv
10	49	35.5	140	CS18_MOUSE	P24460 mus musculu
11	48	35.5	145	CS18_MOUSE	O76096 homo sapien
12	48	34.8	111	CS18_MOUSE	P08935 rattus norv
13	48	34.8	139	CS18_MOUSE	P01038 gallus galli
14	48	34.8	146	CS18_MOUSE	P01034 homo sapien
15	48	34.8	146	CS18_MOUSE	O19093 rattus norv
16	48	34.8	148	CS18_MOUSE	O97862 mus musculu
17	47	34.1	146	CS18_MOUSE	P57351 rattus norv
18	47	34.1	146	CS18_MOUSE	P41555 rattus norv
19	46	33.3	142	CS18_MOUSE	P08935 rattus norv
20	46	33.3	144	CS18_MOUSE	O88098 mus musculu
21	46	33.3	144	CS18_MOUSE	P01038 gallus galli
22	46	33.3	144	CS18_MOUSE	P01034 homo sapien
23	46	33.3	144	CS18_MOUSE	O19093 rattus norv
24	46	33.3	144	CS18_MOUSE	O97862 mus musculu
25	46	33.3	144	CS18_MOUSE	P57351 rattus norv
26	46	33.3	144	CS18_MOUSE	P41555 rattus norv
27	45.5	33.0	652	CS18_MOUSE	P59817 homo sapien
28	45.5	33.0	652	CS18_MOUSE	P59817 homo sapien
29	45	32.6	455	CS18_MOUSE	P59817 homo sapien
30	45	32.6	455	CS18_MOUSE	P59817 homo sapien
31	45	32.6	455	CS18_MOUSE	P59817 homo sapien
32	45	32.6	455	CS18_MOUSE	P59817 homo sapien
33	44.5	32.2	107	CS18_MOUSE	O97563 bos taurus

34	44.5	32.2	170	1	TIM4_RABIT	O97591 oryctolagus
35	44.5	32.2	224	1	TIM4_HUMAN	O99727 homo sapien
36	44.5	32.2	543	1	TIM4_HUMAN	O86972 homo sapien
37	44.5	32.2	1227	1	PR16_HUMAN	O92620 homo sapien
38	44	31.9	436	1	KNL1_BOVIN	P01046 bos taurus
39	44	31.9	447	1	MURE_HELP1	O92366 helicobacte
40	44	31.9	497	1	MURE_HELP1	P57316 buchnera ap
41	44	31.9	621	1	KNH1_BOVIN	P01044 bos taurus
42	44	31.9	3670	1	CSM3_HUMAN	O72407 homo sapien
43	43.5	31.5	1131	1	MOG1_CASEL	P34458 caenorhabdi
44	43	31.2	325	1	FIBP_ADELP	P34774 human adeno
45	43	31.2	459	1	PEL3_ARATH	O9m962 arabidopsis

## ALIGNMENTS

RESULT 1  
ID CS11\_HUMAN STANDARD; PRT; 137 AA.  
AC O9H112; O9H113; 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cysteatin 11 precursor.  
GN CS11 OR CS18.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOPFORMS 1 AND 2).  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.B., Cobby N.R., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Leivaesalho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., McInnes K., McMurtry A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramagay H., Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S., Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sultana J.B., Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;  
"The DNA sequence and comparative analysis of human chromosome 20.";  
Nature 414:665-871(2001).  
-1- SUBCELLULAR LOCATION: Secreted (Potential).  
-1- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Name=1;  
IsoId=O9H112-1; Sequence=Displayed;  
Name=2;  
IsoId=O9H112-2; Sequence=VSP\_001260;  
Note=No experimental confirmation available;  
-1- SIMILARITY: Belongs to the cysteatin family.  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

CC -----  
 DR EMBL; AL096677; CAC13170.1; -  
 DR EMBL; AL096677; CAC17423.1; -  
 DR HSSP; P01038; 1A90.  
 DR Genew; HGNC:15959; CST11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin.1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
 KM Thiol protease inhibitor; Signal; Alternative splicing.  
 FT SIGNAL 1 25  
 FT CHAIN 26 137  
 FT SITE 75 79  
 FT DISULFID 93 101  
 FT DISULFID 114 134  
 FT CARBOHYD 131 131  
 FT VARSPIC 76 110  
 FT SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;  
 Query Match 100.0%; Score 138; DB 1; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-14;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHILNVEMQWTCCKPE 24  
 Db 74 RQVTDHLEHILNVEMQWTCCKPE 97

RESULT 2  
 CS11\_MOUSE STANDARD; PRT; 139 AA.  
 ID CS11\_MOUSE  
 AC Q9D269;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin 11 precursor.  
 GN CST11.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kanukawa T., Saito R.,  
 RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Peele G., Quackenbush J.,  
 RA Schirrl L. M., Seabull F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Botfield D., Boujarda N., Carninci P., de Bonaldo M. F.,  
 RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gusticich S., Hill D., Hofmann M., Hume D. A., Kamita M., Lee N. H.,  
 RA Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombereis P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakanoto N.,  
 RA Sasakura H., Sato K., Schoenbach C., Seta Y., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyokawa K., Wang K. H., Weitz C., Whitaker C., Wilting L.,  
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

CC -----  
 DR EMBL; AK020300; BAB32061.1; -  
 DR HSSP; P01034; 1G96.  
 DR MED; MG1:1925490; Cst11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin.1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE NEG.  
 KM Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 28  
 FT CHAIN 29 139  
 FT SITE 76 80  
 FT DISULFID 94 102  
 FT DISULFID 115 135  
 FT CARBOHYD 134 134  
 FT SEQUENCE 139 AA; 16217 MW; F228D9815FA32640 CRC64;  
 Query Match 65.9%; Score 91; DB 1; Length 139;  
 Best Local Similarity 70.8%; Pred. No. 8.5e-07;  
 Matches 17; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RQVTDHLEHILNVEMQWTCCKPE 24  
 Db 75 RQVTDHLEHILNVEMQWTCCKPE 98

RESULT 3  
 CST8\_RAT STANDARD; PRT; 142 AA.  
 ID CST8\_RAT  
 AC O88969;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin  
 DE B).  
 GN CST8 OR CRES.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Epididymis;  
 RX MEDLINE=99247899; PubMed=10229662;  
 RA Cornwall G. A., Hsia N., Sutton H. G.;  
 RT "Structure, alternative splicing and chromosomal localization of the  
 RL Cystatin-related epididymal spermatogenic gene."  
 RL Biochem. J. 340:85-93(1999).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

CC -----  
 DR EMBL; AF090692; AAC6317.1; -  
 DR HSSP; P01034; 1G96.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin.1.  
 DR SMART; SM00043; CY; 1.

KW Thiol protease inhibitor; Signal.  
 PT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 FT SITE 77 81 PROTEIN.  
 FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 119 139 BY SIMILARITY.  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 142 AA; 16246 MW; PB873FAA6BCAB34 CRC64;  
 Query Match 51.4%; Score 71; DB 1; Length 142;  
 Best Local Similarity 45.5%; Pred. No. 0.00092;  
 Matches 10; Conservative 8; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 QVTDHLEHYLVEMQWTCCKP 23  
 77 QITDRMEHYDVOISRSNCKRP 98  
 Db  
 RESULT 4  
 CST8 MOUSE STANDARD; PRT; 142 AA.  
 ID CST8\_MOUSE  
 AC P32766; 089102;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermato-genic protein precursor (Cystatin-  
 related epididymal specific protein) (Cystatin 8).  
 GN CST8 OR CRGS.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H, and CD-1;  
 RX MEDLINE=99247899; PubMed=10229662;  
 RA Cornwall G.A., Hsia N., Sutton H.G.,  
 RT "Structure, alternative splicing and chromosomal localization of the  
 RT cystatin-related epididymal spermato-genic gene.";  
 RL Biochem. J. 340:85-93(1999).  
 RN [2]  
 RP SEQUENCE OF 4-142 FROM N.A.  
 RC TISSUE=Epididymis;  
 RX MEDLINE=93078799; PubMed=1280328;  
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.,  
 RT "The CRGS gene: a unique testis-regulated gene related to the cystatin  
 RT family is highly restricted in its expression to the proximal region  
 RT of the mouse epididymis.";  
 RL Mol. Endocrinol. 6:1653-1664(1992).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epididymium.  
 CC -1- INDUCTION: Testicular factors or hormones other than androgens  
 CC present in the testicular fluid may be involved in the regulation  
 CC of CRGS gene expression.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; AF091503; AAC61754.1; -  
 CC DR EMBL; AF090691; AAC6316.1; -  
 CC EMBL; S49926; AAC35390.1; -

DR PIR; A45361; A45361.  
 DR HSSP; P01034; 1G96.  
 DR MGD; MGI:107161; Cst8.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 KW Thiol protease inhibitor; Signal.  
 PT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 FT SITE 77 81 PROTEIN.  
 FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 119 139 BY SIMILARITY.  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 4 15 PMLWSLFLFTIP -> GTRDQVGSQK (IN REF. 2).  
 SQ SEQUENCE 142 AA; 16288 MW; 50B446B98F6673E CRC64;  
 Query Match 45.7%; Score 63; DB 1; Length 142;  
 Best Local Similarity 40.9%; Pred. No. 0.015;  
 Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 QVTDHLEHYLVEMQWTCCKP 23  
 77 QITDRMEHYDVOISRSNCKRP 98  
 Db  
 RESULT 5  
 CST8 HUMAN STANDARD; PRT; 142 AA.  
 ID CST8\_HUMAN  
 AC O60676;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermato-genic protein precursor (Cystatin  
 related epididymal specific protein) (Cystatin 8).  
 GN CST8 OR CRGS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=95344753; PubMed=7619504;  
 RA Cornwall G.A., Hann S.R.,  
 RT "Transient appearance of CRGS protein during spermatogenesis and  
 RT caput epididymal sperm maturation.";  
 RL Mol. Reprod. Dev. 41:37-46(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=1638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagunley C.L.,  
 RA Bailey W., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Gardner C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhand P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leveson-Smith M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McEay K., McMurtry A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.P., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
 RA Rice C.M., Rose M.T., Scott C.E., Sehara H.K., Showkseen R., Sims S.,  
 RA Spruce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,



```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Buffalo;
RX MEDLINE=90092122; PubMed=2689174;
RA Cole T., Dickson P.W., Esnard F., Averill F., Risbridger G.,
RT Gauthier F., Schneider G.;
RA "The cDNA structure and expression analysis of the genes for the
RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
RL in rat brain.";
RL Eur. J. Biochem. 186:35-42(1989).
RN [2]
RP SEQUENCE OF 8-127.
RX MEDLINE=90380276; PubMed=2400577;
RA Esnard F., Esnard A., Faucher D., Capony J.-P., Derancourt J.,
RA Brillard M., Gauthier F.;
RT "Rat cystatin C: the complete amino acid sequence reveals a site for
RL N-glycosylation.";
RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).
RN [3]
RP SEQUENCE OF 8-49.
RX MEDLINE=88313020; PubMed=3044831;
RA Esnard A., Esnard F., Faucher D., Gauthier F.;
RT "Two rat homologues of human cystatin C.";
RL FEBS Lett. 236:475-478(1988).
RN [4]
RP SEQUENCE OF 8-20.
RC TISSUE=Sertoli cells;
RX MEDLINE=92225121; PubMed=1563513;
RA Esnard A., Esnard F., Guillou F., Gauthier F.;
RT "Production of the cysteine proteinase inhibitor cystatin C by rat
RT Sertoli cells.";
RL FEBS Lett. 300:131-135(1992).
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity. Known to inhibits cathepsin B,
CC H, and L.
CC
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16957; CAA34831.1; -.
DR PIR; S07085; S07085.
DR PIR; S10587; S10587.
DR HSSP; P01034; 1G96.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
DR Thiol protease inhibitor; Signal.
FT NON TER 1 1
FT SIGNAL <1 7
FT CHAIN 8 127
FT ACT SITE 18 18 CYSTATIN C.
FT SITE 62 66 REACTIVE SITE.
FT DISULFID 80 90 SECONDARY AREA OF CONTACT.
FT DISULFID 104 124 BY SIMILARITY.
FT CONFLICT 25 25 A -> E (IN REF. 2).
SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;
Query Match 38.4%; Score 53; DB 1; Length 127;
Best Local Similarity 41.7%; Pred. No. 0.43;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0

```

[illegible]

```

AC 019093;
DT 15-JUN-1998 (Rel. 36, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN Cst3.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C, Icelandic-like mutation in an animal model of
RT cerebrovascular beta-amyloidosis.";
RL Stroke 27:2080-2085(1996).
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ibt-sib.ch/announce/
CC or send an email to license@ibt-sib.ch).
CC -----
DR EMBL; U52028; AAB64051.1; -.
DR HSSP; P01034; 1G96.
DR InterPro; IPR00010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; Cy.1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 146 CYSTATIN C.
FT ACT SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 15946 MW; 08196353C0306AA3 CRC64;
Query March 36.2%; Score 50; DB 1; Length 146;
Best Local Similarity 45.5%; Pred. No. 1.4;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 1 ROVTDLHLYHLNVEWMTTCOK 22
Db 80 KQIVAGVNYFLDVMGRITTC 101
RESULT 10
CYTC MOUSE STANDARD; PRT; 140 AA.
ID CYTC MOUSE
AC P21460;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin C precursor (Cystatin 3).
GN Cst3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=91054522; PubMed=2241983;
RA Solem W., Rawson C., Lindburg K., Barnes D.;
RA "Transforming growth factor beta regulates cystatin C in serum-free

```

```

RT mouse embryo (SFME) cells.";
RN Biochem. Biophys. Res. Commun. 172:945-951(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=129/Sv; TISSUE=Liver;
RX MEDLINE=95137392; PubMed=7835704;
RA Hub C., Nagle J.W., Kozak C.A., Abrahamson M., Karlsson S.;
RA "Structural organization, expression and chromosomal mapping of the
RT mouse cystatin-C-encoding gene (Cst3).";
RL Gene 152:221-226 (1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=ILS, and ILS;
RX MEDLINE=21363810; PubMed=11471062;
RA Ehninger M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikelia J.M.;
RT "High-throughput sequence identification of gene coding variants
RT within alcohol-related QTLs.";
RL Mamm. Genome 12:657-663(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marisla K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uudin T.B., Toshiyuki S., Cantanci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Hellon E., Keltman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Butcherfield A.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ibt-sib.ch/announce/
CC or send an email to license@ibt-sib.ch).
CC -----
DR EMBL; M59470; AAB63298.1; -.
DR EMBL; U10098; AAB41056.1; -.
DR EMBL; AF483486; AAL90760.1; -.
DR EMBL; AF483487; AAL90761.1; -.
DR EMBL; BC002072; AAH02072.1; -.
DR PIR; A36163; A36163.
DR HSSP; P01034; 1G96.
DR MGD; MGI:102519; Cst3.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 20
FT CHAIN 21 140 CYSTATIN C.
FT ACT SITE 31 31 REACTIVE SITE.
FT SITE 75 79 SECONDARY AREA OF CONTACT.
FT DISULFID 93 103 BY SIMILARITY.

```

FT DISURFID 117 137 BY SIMILARITY.  
 FT CONFLICT 16 16 A -> G (IN REF. 1).  
 FT CONFLICT 84 84 L -> F (IN REF. 1).  
 SQ SEQUENCE 140 AA; 15531 MW; 3A563406D58D0F5 CRC64;  
 Query March 35.5%; Score 49; DB 1; Length 140;  
 Best Local Similarity 41.7%; Pred. No. 1.9;  
 Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 RQVTHLEHYNVEMQWTCCKPE 24  
 Db 74 KQVAGVNFVLDVEMGRCTCKSQ 97  
 RESULT 11  
 CYTE\_HUMAN STANDARD; PRT; 145 AA.  
 ID 076096; O9UED4;  
 AC 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin F precursor (leukocystatin) (Cystatin 7) (Cystatin-like  
 DE metacystatin-associated protein) (CMAP).  
 GN CST7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98406133; PubMed=9733783;  
 RA Ni J., Fernandez M.A., Danielsson L., Chiklaskuru R.A., Zhang J.,  
 RA Grubb A., Su J., Gentz R., Abrahamson M.,  
 RT "Cystatin F is a glycosylated human low molecular weight cysteine  
 RT proteinase inhibitor."  
 RT J. Biol. Chem. 273:24797-24804 (1998).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9828157; PubMed=9632704;  
 RA Halton S., Ford J., Foster J., Dowling L., Lucian L., Steerling M.,  
 RA Xu Y., Weiss M., Ikeda M., Liggett D., Helms A., Caux C., Lebecque S.,  
 RA Hannum C., Menon S., McClanahan T., Gorman D., Zurawski G.,  
 RT "leukocystatin, a new class II cystatin expressed selectively by  
 RT hematopoietic cells."  
 RT J. Biol. Chem. 273:16400-16408 (1998).  
 RL [3]  
 RP SEQUENCE FROM N.A.  
 RA Morita M., Arakawa H., Yoshiuchi N.,  
 RT "Human homologue of murine CMAP."  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20399571; PubMed=10945474;  
 RA Morita M., Hara Y., Tamai Y., Arakawa H., Nishimura S.,  
 RT "Genomic construct and mapping of the gene for CMAP  
 RT (leukocystatin/Cystatin F, CST7) and identification of a proximal  
 RT novel gene, BSCV (C20orf3)."  
 RT Genomics 67:87-91 (2000).  
 RL [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stevrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lhvaeslahti M.H., Leyerle M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McEay K., McMurtry A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
 RA Rice C.M., Ross M.L., Scott C.E., Sehra H.K., Showkseen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sultson J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams J., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.,  
 RT "The DNA sequence and comparative analysis of human chromosome 20."  
 RL Nature 414:865-871 (2001).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uudin T.B., Yoshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,  
 RA Bosa S.A., McEwan P.J., McKernan K.J., Malek A.M., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.C., Krzywinski M.I., Skalka U., Smalins D.E.,  
 RA Schercher A., Schein J.B., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- FUNCTION: Inhibits pepsin and cathepsin B but with affinities  
 CC lower than other cystatins. May play a role in immune regulation  
 CC through inhibition of a unique target in the hematopoietic system.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- TISSUE SPECIFICITY: Primarily expressed in peripheral blood cells  
 CC and spleen.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1ab-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL, AF036342; AAC35747.1; -;  
 DR EMBL, AF031824; AAC39788.1; -;  
 DR EMBL, AB015225; BAA34941.1; ALT\_INIT.  
 DR EMBL, AB029636; BAB11886.1; ALT\_INIT.  
 DR EMBL, AL035661; CAB75498.1; -;  
 DR EMBL, BC015507; AAH15507.1; ALT\_INIT.  
 DR HSSP, P01034; 1G96.  
 DR Genew, HGNC:2479; CST7.  
 DR MIM, 603253; -;  
 DR GO, GO:0004869; F:cysteine protease inhibitor activity; TAS.  
 DR GO, GO:0006955; P:immune response; TAS.  
 DR InterPro, IPR000010; Cystatin.  
 DR Pfam, PF00003; cystatin; 1.  
 DR SMART, SM00043; Cy; 1.  
 DR PROSITE, PS00287; CYSTATIN, FALSE\_NEG.  
 DR Thiol protease inhibitor; Glycoprotein; signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 145 CYSTATIN F.

```

FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 110 BY SIMILARITY.
FT DISULFID 124 144 BY SIMILARITY.
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 145 AA; 16454 MW; B2BCA4F6857CB0F CRC64;

Query Match 35.5%; Score 49; DB 1; Length 145;
Best Local Similarity 47.6%; Pred. No. 2;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 QVTDHLEVHLNVEWQMTTCOK 22
Db 81 QIVVGLKXMLEVEIGRTTCKK 101

RESULT 12
CYT_BITAR STANDARD; PRT; 111 AA.
ID_CYT_BITAR
AC P08935;
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin.
OS Bitis arietans (African puff adder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
OC Viperidae; Viperinae; Bitis.
OX NCBI_TaxID=8692;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=88076861; PubMed=3500714;
RA Ritonja A., Evans H.J., Machleidt W., Barrett A.J.;
RT "Amino acid sequence of a cystatin from venom of the African puff
RT adder (Bitis arietans).";
RL Biochem. J. 246:799-802(1987).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the cystatin family.
DR HSSP; A28793; A28793.
DR HSSP; P01038; ICW.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KM Thiol protease inhibitor.
FT ACT_SITE 3 3 REACTIVE SITE.
FT SITE 47 51 SECONDARY AREA OF CONTACT.
FT DISULFID 65 81 PROBABLE.
FT VARIANT 57 57 M -> T (IN EQUAL AMOUNT).
SQ SEQUENCE 111 AA; 12678 MW; 1A31B6B246AC10C CRC64;

Query Match 34.8%; Score 48; DB 1; Length 111;
Best Local Similarity 42.9%; Pred. No. 2.1;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEVHLNVEWQMTTCOK 22
Db 47 QVVGSKRYLMELKTTCKK 67

RESULT 13
CYT_CHICK STANDARD; PRT; 139 AA.
ID_CYT_CHICK
AC P01036;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin precursor (Egg-white cystatin).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

```

```

OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90008673; PubMed=2793849;
RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;
RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,
RT and tissue distribution.";
RL J. Biol. Chem. 264:17164-17169(1989).
RN [2]
RP SEQUENCE OF 24-139
RX MEDLINE=84178305; PubMed=6712597;
RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;
RT "Cystatin. Amino acid sequence and possible secondary structure.";
RL Biochem. J. 217:813-817(1984).
RN [3]
RP SEQUENCE OF 24-139
RX MEDLINE=84110059; PubMed=6662498;
RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,
RA Machleidt W.;
RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
RT of cystatin from chicken egg white.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
RN [4]
RP CHARACTERIZATION OF PROTEIN.
RX MEDLINE=83256421; PubMed=6409085;
RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,
RA Sunter D.C., Barrett A.J.;
RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved
RT purification from egg white, characterization, and detection in
RT chicken serum.";
RL Biochem. J. 211:129-138(1983).
RN [5]
RP DISULFIDE BONDS.
RX Grubb A., Loeberg H., Barrett A.J.;
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
RT cystatin.";
RL FEBS Lett. 170:370-374(1984).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=89252033; PubMed=2721673;
RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,
RA Bode W.;
RT "The cysteine proteinase inhibitor chicken cystatin is a
RT phosphoprotein.";
RL FEBS Lett. 248:162-168(1989).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=89052676; PubMed=3191914;
RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshkov A.,
RA Brzin J., Kos J., Turk V.;
RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and
RT its possible mode of interaction with cysteine proteinases.";
RL EMBO J. 7:2593-2599(1988).
RN [8]
RP STRUCTURE BY NMR.
RX MEDLINE=94087719; PubMed=8263912;
RA Dieckmann T., Mitschng U., Hofmann M., Kos J., Turk V.,
RA Auerwald E.A., Jeanicke R., Oeschkinat H.;
RT "The structures of native phosphorylated chicken cystatin and of a
RT recombinant unphosphorylated variant in solution.";
RL J. Mol. Biol. 234:1048-1059(1993).
CC -1- FUNCTION: This protein binds tightly to and inhibits a variety of
CC thiol proteases including ficin, papain, and cathepsins B, C, H,
CC and L. Although isolated from egg white, it is also present in
CC serum.
CC -1- SIMILARITY: Belongs to the cystatin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL, J05077, AAA48744.1; -.

DR PIR, A34456; UDCH.

DR PDB, 1CEW, 31-JAN-94.

DR PDB, 1A67, 17-MAY-98.

DR PDB, 1A90, 17-JUN-98.

DR InterPro: IPR000010, Cystatin.

DR Pfam: PF00031, Cystatin, 1.

DR SMART, SM00043, CY, 1.

DR PROSITE, PS00287, CYSTATIN, 1.

KM Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.

FT SIGNAL 1 23

FT CHAIN 24 139

FT ACT\_SITE 32 32

FT SITE 76 80

FT DISULFID 94 104

FT DISULFID 118 138

FT MOD\_RES 103 103

FT STRAND 35 36

FT TURN 39 40

FT TURN 42 51

FT TURN 52 52

FT HELIX 53 56

FT TURN 57 58

FT STRAND 63 77

FT TURN 81 95

FT TURN 96 97

FT TURN 99 100

FT HELIX 101 108

FT STRAND 115 125

FT TURN 126 129

FT STRAND 130 139

SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;

Query March 34.8%; Score 48; DB 1; Length 139;

Best Local Similarity 45.5%; Pred. No. 2.7;

Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 ROYTDHLEHNLVEMQWTTCK 22

DB 75 RQLVSGIKYILQVEIGRTTCK 96

RESULT 14

CYTC HUMAN

AC CYTC HUMAN STANDARD; PRT; 146 AA.

AC P01034;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)

DE (Post-gamma-globulin).

GN Ctr3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OC NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RA Abrahamsen M., Grubb A., Olafsson I., Lundwall A.;

RT "Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C.";

RL FEBS Lett. 216:229-233(1987).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Leukocyte;

RA Abrahamsen M., Olafsson I., Paledottir A., Ulvbaeck M., Lundwall A.,

RA Jensen O., Grubb A.;

RT "Structure and expression of the human cystatin C gene.";

RL Biochem. J. 268:287-294(1990).

RN [3]

RP SEQUENCE FROM N.A. (HCMA VARIANT).

RC TISSUE=Brain;

RA MEDLINE=69235594; PubMed=2541223;

RA Levy E., Lopez-Otin C., Ghiso J., Gellner D., Frangione B.;

RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a mutation in the cystatin C gene, an inhibitor of cysteine proteases.";

RT J. Exp. Med. 169:1771-1778(1989).

RL [4]

RP SEQUENCE FROM N.A.

RA MEDLINE=69350949; PubMed=2764935;

RA Satoh E., Sabatini L.M., Eddy R.L., Snows T.B., Azen E.A.,

RA Isemura S., Sanada K.;

RT "The human cystatin C gene (CST3) is a member of the cystatin gene family which is localized on chromosome 20.";

RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).

RN [5]

RP SEQUENCE FROM N.A.

RA Dickinson D.P., Hewett-Emmett D., Thiesse M.;

RT "Acquisition of complex patterns of differential expression in epithelial cell populations during the evolution of type 2 cystatin genes.";

RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

RL [6]

RP SEQUENCE FROM N.A.

RA MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,

RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,

RA Beasley O.P., Bird C.F., Blakey S.E., Bridgman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cobby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Griffin D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA Kay M.P., Kimberley A.M., King A., Knights K., Layrd G.K., Lawlor S.,

RA Lehaealho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConachie I.J., McElay K., McWherry A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,

RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,

RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,

RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sultson J.E.,

RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W.,

RA Whitehead S.L., Whitaker P., Willey D.L., Williams S.A.,

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

RA Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.";

RL Nature 414:865-871(2001).

RN [7]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA MEDLINE=22388257; PubMed=12477932;

RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Kluemper R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Blat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,

RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,

RA Rana S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton B., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences."; *Proc. Natl. Acad. Sci. U.S.A.* 99:16899-16903(2002).  
 RN [8]  
 RN SEQUENCE OF 27-146.  
 RX MEDLINE=8222268; PubMed=6283552;  
 RA Grubb A., Loeferberg H.;  
 RT "Human gamma-trace, a basic microprotein: amino acid sequence and  
 presence in the adenohypophysis."; *Proc. Natl. Acad. Sci. U.S.A.* 79:3024-3027(1982).  
 RN [9]  
 RN SEQUENCE OF 27-73.  
 RX MEDLINE=84110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,  
 RA Machleidt W.;  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 of cystatin from chicken egg white."; *Hoppe-Seyler's Z. Physiol. Chem.* 364:1487-1496(1983).  
 RN [10]  
 RN SEQUENCE OF 27-76.  
 RX MEDLINE=84128015; PubMed=6365094;  
 RA Brzin J., Popovic T., Turk V.;  
 RT "Human cystatin, a new protein inhibitor of cysteine proteinases."; *Biochem. Biophys. Res. Commun.* 118:103-109(1984).  
 RN [11]  
 RN DISULFIDE BONDS.  
 RA Grubb A., Loeferberg H., Barrett A.J.;  
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
 cystatin."; *FEBS Lett.* 170:370-374(1984).  
 RN [12]  
 RN X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.  
 RX MEDLINE=21173909; PubMed=11276250;  
 RA Janowski R., Kozak M., Jankowska E., Grzonka Z., Grubb A.,  
 RA Abrahamson M., Jaskolski M.;  
 RT "Human cystatin C, an amyloidogenic protein, dimerizes through  
 three-dimensional domain swapping."; *Nat. Struct. Biol.* 8:316-320(2001).  
 RN [13]  
 RN VARIANT GLN-94.  
 RX MEDLINE=92316504; PubMed=1352269;  
 RA Abrahamson M., Jonedottir S., Olafsson I., Jansson O., Grubb A.;  
 RT "Hereditary cystatin C amyloid angiopathy: identification of the  
 disease-causing mutation and specific diagnosis by polymerase chain  
 reaction based analysis."; *Hum. Genet.* 89:377-380(1992).  
 RL Hum. Genet. 89:377-380(1992).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 thought to serve an important physiological role as a local  
 regulator of this enzyme activity.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,  
 vas deferens, brain, thymus, and ovary and the lowest in the  
 submandibular gland.  
 CC -1- DISEASE: Defects in CST3 are a cause of hereditary cerebral  
 hemorrhage with amyloidosis (HCHWA) [MM:105150]; also known as  
 cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis  
 Icelandic type. HCHWA is characterized by a thickening of the  
 cerebral arteries walls with deposition of material with the  
 characteristics of amyloid.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X05607; CAA29096.1; -;  
 DR EMBL; X52255; CAA36497.1; -;  
 DR EMBL; M27891; AAA52164.1; -;  
 DR EMBL; M27889; AAA52164.1; JOINED.  
 DR EMBL; M27890; AAA52164.1; JOINED.  
 DR EMBL; X61681; CAA43856.2; -;  
 DR EMBL; X61682; CAA43856.2; JOINED.  
 DR EMBL; AF319564; AAK11570.1; -;  
 DR EMBL; AL121894; CAC05424.1; -;  
 DR EMBL; BC013083; AAH13083.1; -;  
 DR PIR; S10216; UDHU.  
 DR PDB; 1G96; 06-APR-01.  
 DR Genew; HGNC:2475; CST3.  
 DR MIM; 604312; -;  
 DR MIM; 105150; -;  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 KW Thiol protease inhibitor; Amyloid; Signal; Disease mutation;  
 KW Polymorphism; 3D-structure.  
 FT SIGNAL 1 26  
 FT CHAIN 27 146  
 FT ACT SITE 37 37  
 FT SITE 81 85  
 FT DISULFD 99 109  
 FT DISULFD 123 143  
 FT  
 Query Match 34.8%; Score 48; DB 1; Length 146;  
 Best Local Similarity 37.5%; Pred. No. 2.8;  
 Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 RQVTDHLEFHLNVEMQWTCQKE 24  
 DB 80 KQIVAGVYFLDVELGRCTCTKTQ 103  
 ID CYTC\_MACMU STANDARD; PRT; 146 AA.  
 AC O19092;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin C precursor.  
 GN CST3.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 CC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=97054523; PubMed=8898820;  
 RA Wei L.H., Walker L.C., Levy E.;  
 RT "Cystatin C, Icelandic-like mutation in an animal model of  
 cerebrovascular beta-amyloidosis."; *Stroke* 27:2080-2085(1996).  
 RL Stroke 27:2080-2085(1996).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 thought to serve an important physiological role as a local  
 regulator of this enzyme activity.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U51912; AAB64050.1; -;

Fri Mar 19 07:21:55 2004

DR HSSP: P01034; 1G96.  
DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; Cystatin; 1.  
DR SMART: SM00043; CY; 1.  
DR PROSITE: PS00287; CYSTATIN; 1.  
KM Thiol protease inhibitor; Amyloid; Signal.  
FT SIGNAL 1 26 BY SIMILARITY.  
FT CHAIN 27 146 CYSTATIN C.  
FT ACT\_SITE 37 37 REACTIVE SITE.  
FT SITE 81 85 SECONDARY AREA OF CONTACT.  
FT DISULFID 99 109 BY SIMILARITY.  
FT DISULFID 123 143 BY SIMILARITY.  
SQ SEQUENCE 146 AA; 15857 MW; F083B8774A29DF26 CRC64;  
  
Query Match 34.8%; Score 48; DB 1; Length 146;  
Best Local Similarity 37.5%; Pred. No. 2.8;  
Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
  
QY 1 ROYTDHLEVHLNEMQWTCOKPE 24  
DB 80 KQIVAGVNFPLDVELGRITCKTQ 103

Search completed: March 18, 2004, 14:16:12  
Job time : 3.85714 secs





DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)  
 DE Cystatin 11.  
 GN CST11.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RA Hamil K.G., Hall S.H.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF501290; AAM21709.1; -  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00003; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 SQ SEQUENCE 139 AA; 16686 MW; E1E36DB786B4D08C CRC64;

Query Match 60.9%; Score 84; DB 11; Length 139;  
 Best Local Similarity 58.3%; Pred. No. 2.2e-05;  
 Matches 14; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 QVTDHLEHYLVNEMQWTTCKP 24  
 DB 75 KQMTNMEFHITVEMORTTCKTE 98

## RESULT 3

Q9DAP1 PRELIMINARY; PRT; 141 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)

DE 1700006C19R1k protein.

GN 1700006C19R1k.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=CS7BL/6J; TISSUE=Testis;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinsagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,

RA Aizawa K., Itawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita C., Whitaker C., Wilting L.,

RA Guelincich S., Hill D., Hofmann M., Mazarelli J., Mombereis P.,

RA Lyons P., Marchionni L., Mashima J., Mazarelli J., Mombereis P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyo-oka K., Wang K.H., Weiss C., Whitaker C., Wilting L.,

RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayaishi Y.,

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AK005665; BAB24175.1; -

DR HSSP; P01038; ICEW.

DR MGD; MGI:196544; 1700006C19R1k.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro: IPR000010; Cystatin.

DR Pfam: PF00003; Cystatin; 1.

DR SMART; SM00043; CY; 1.

SQ SEQUENCE 141 AA; 16811 MW; C20FA0DB8B1AC378C CRC64;

Query Match 43.5%; Score 60; DB 11; Length 141;  
 Best Local Similarity 52.4%; Pred. No. 0.14;  
 Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEHYLVNEMQWTTCK 22  
 DB 76 QITDSLEYLVNARTMCKK 96

## RESULT 4

Q80ZNS PRELIMINARY; PRT; 141 AA.

DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, last sequence update)

DT 01-JUN-2003 (TREMBlrel. 25, last annotation update)

DE RIKEN cDNA 1700006C19 gene.

GN Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Testicle;

RA Straube R.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC048681; AAH48681.1; -

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro: IPR003243; Cystatin.

DR Pfam; PF00031; Cystatin; 1.

DR Prodom; PD001231; Cystatin\_C/M; 1.

DR SMART; SM00043; CY; 1.

SQ SEQUENCE 141 AA; 16825 MW; C20FA0DB8B1AC378C CRC64;

Query Match 43.5%; Score 60; DB 11; Length 141;  
 Best Local Similarity 52.4%; Pred. No. 0.14;  
 Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEHYLVNEMQWTTCK 22  
 DB 76 QITDSLEYLVNARTMCKK 96

## RESULT 5

Q8DS42 PRELIMINARY; PRT; 236 AA.

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)

DE Putative ABC transporter, ATP-binding protein.

GN ADC OR SMU.1994.

OS Streptococcus mutans.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=1309;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=UA159 / ATCC 700610 / Serotype C;

RX MEDLINE=22295063; PubMed=12397186;

RA Ajdic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,

RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Ikin S., Qian Y.,

RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferrerelli J.J.;

RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental

pathogen.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).

DR EMBL; AE015022; AAN59598.1; -

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.

DR GO; GO:0000166; F:nucleotide binding; IEA.

GO: 0006810; P:transport; IEA.  
 DR InterPro: IPR003593; AAA\_Atpase.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR Pfam: PF00005; ABC\_tran; 1.  
 DR ProDom: PD000006; ABC\_transporter; 1.  
 DR SMART: SM00382; AAA; 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; 1.  
 DR PROSITE: PS50893; ABC\_TRANSPORTER\_2; 1.  
 KM ATP-binding; Complete proteome.  
 SQ SEQUENCE 236 AA; 26812 MW; 1BAB8DD742DC76DB CRC64;  
 Query Match 41.3%; Score 57; DB 16; Length 236;  
 Best Local Similarity 37.0%; Pred. No. 0.7;  
 Matches 10; Conservative 6; Mismatches 7; Indels 4; Gaps 1;  
 QY 1 RQVTHLEHNLVEMQ-----WTCQKP 23  
 Db 108 RRLTKHDDHKKVLESGVMWENCQKP 134  
 RESULT 6  
 Q94M04 PRELIMINARY; PRT; 723 AA.  
 ID 094M04  
 AC 094M04  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Major core structural protein pI.  
 GN 1.  
 OS Bacteriophage phi-12.  
 OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.  
 NC NCBI\_TaxID=161736;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Gottlieb P.J., Potgieter C., Wei H., Toporovsky I.;  
 RT "Characterization of Bacteriophage phi12."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF086836; AL010109.1; -  
 DR InterPro: IPR006162; Ppanthe.S.  
 DR PROSITE: PS00012; PHOSPHOPANTHINE; 1.  
 SQ SEQUENCE 723 AA; 78950 MW; F5DEF7FF93A50B7C CRC64;  
 Query Match 41.3%; Score 57; DB 9; Length 723;  
 Best Local Similarity 50.0%; Pred. No. 2.3;  
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
 QY 6 HLEHNLVEMQWTCQKP 23  
 Db 463 HTEYNRFVDLQWATCYP 480  
 RESULT 7  
 018150 PRELIMINARY; PRT; 1779 AA.  
 ID 018150  
 AC 018150; O45627;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE T28B8.4 protein.  
 GN T28B8.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;  
 OC Rhabditidae; Peioderinae; Caenorhabditis.  
 NC NCBI\_TaxID=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA White S.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN (12)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkes T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kerhaw J., Kisten T., Laister N., Lattelle P.,  
 RA Lightning U., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifen L., Roopra A., Saunders D., Showkeen R.,  
 RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sprout J., Woldman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38 (1994).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RA White S.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z81133; CAB03445.1; -  
 DR EMBL; AL021066; CAB03445.1; JOINED.  
 DR EMBL; AL021066; CAI5925.1; -  
 DR EMBL; Z81133; CAI5925.1; JOINED.  
 DR PIR; T23130; T23130.  
 DR WormPep; T28B8.4; CBI6519.  
 DR InterPro: IPR008938; ANM.  
 SQ SEQUENCE 1779 AA; 205511 MW; 70A4489148B96F3B CRC64;  
 Query Match 38.4%; Score 53; DB 5; Length 1779;  
 Best Local Similarity 37.0%; Pred. No. 26;  
 Matches 10; Conservative 5; Mismatches 4; Indels 8; Gaps 1;  
 QY 1 RQVTHLEHNLVEMQ-----WTCQKP 19  
 Db 749 RRLTKHDDHKKVLESGVMWENCQKP 775  
 RESULT 8  
 08T9G8 PRELIMINARY; PRT; 540 AA.  
 ID 08T9G8  
 AC 08T9G8  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE SD01519P.  
 GN AATS-ALA OR CG13391.  
 GN Drosophila melanogaster (fruit fly).  
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygotia;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCBI\_TaxID=7227;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY069773; AL33918.1; -  
 DR FLYbase; FBgn027094; Aats-ala.  
 DR GO; GO:0004813; F:alanine-tryptophan ligase activity; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0006419; P:alanine-tryptophan aminoacylation; IEA.  
 DR InterPro: IPR003156; DHHA1.  
 DR InterPro: IPR002318; CRNA-synt\_2c.  
 DR InterPro: IPR006193; CRNA-synt\_Ala.  
 DR Pfam; PF02272; DHHA1; 1.  
 DR Pfam; PF01411; CRNA-synt\_2c; 1.  
 DR PROSITE: PS50860; AA-TRNA-LIGASE II ALA; 1.  
 SQ SEQUENCE 540 AA; 59580 MW; B670DCFC386843B CRC64;  
 Query Match 37.7%; Score 52; DB 5; Length 540;  
 Best Local Similarity 50.0%; Pred. No. 10;  
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEMQWTT 19  
 Db 157 KVGDELEHLIDVERRWLT 174

## RESULT 9

ID Q8T9K4 PRELIMINARY; PRT; 966 AA.  
 AC Q8T9K4;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE GN03058P.  
 GN AATS-ALA OR CG13391.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxId=7227;  
 RX SEQUENCE FROM N.A.  
 RA Chadapton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Doreet V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Munoz J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.,  
 RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AY069255; AL339400.1; -  
 DR FLYBase; FBgn027094; Aats-ala.  
 DR GO; GO:0004813; F:alanine-cRNA ligase activity; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0006419; F:nucleic acid binding; IEA.  
 DR GO; GO:0006419; P:alanyl-cRNA aminoacylation; IEA.  
 DR InterPro; IPR003156; DHHA1.  
 DR InterPro; IPR002318; cRNA-synt\_2c.  
 DR InterPro; IPR006193; cRNA-synt\_1a.  
 DR Pfam; PF02272; DHHA1; 1.  
 DR Pfam; PF01411; cRNA-synt\_2c; 1.  
 DR PRINTS; PR00980; TRNASYNTHALA.  
 DR TIGRFAMs; TIGR00344; alas; 1.  
 DR PROSITE; PSS0860; AA\_TRNA\_LIGASE\_II\_ALA; 1.  
 DR SEQUENCE 966 AA; 107697 MW; 9F2079A119F463B4 CRC64;  
 SQ  
 Query Match 37.7%; Score 52; DB 5; Length 966;  
 Best Local Similarity 50.0%; Pred. No. 19;  
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 QVTDHLEHYHLNVEMQWTT 19  
 Db 583 KVGDELEHLIDVERRWLT 600  
 RESULT 10  
 ID Q9VLM8 PRELIMINARY; PRT; 966 AA.  
 AC Q9VLM8;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE AATS-ALA protein.  
 GN AATS-ALA OR CG13391.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxId=7227;  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=Berkeley;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Vandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burdick K.C., Buzam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferrer A., Fleischmann W.,  
 RA Fobel C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser R.,  
 RA Glodok A., Gong P., Gorrell J.H., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,  
 RA Jatalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Morkulov G., Misha N.V., Moberly C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard U., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E.E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkovic R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AEO03621; AAT52657.1; -  
 DR FLYBase; FBgn027094; Aats-ala.  
 DR GO; GO:0004813; F:alanine-cRNA ligase activity; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0006419; F:nucleic acid binding; IEA.  
 DR GO; GO:0006419; P:alanyl-cRNA aminoacylation; IEA.  
 DR InterPro; IPR003156; DHHA1.  
 DR InterPro; IPR002318; cRNA-synt\_2c.  
 DR InterPro; IPR006193; cRNA-synt\_1a.  
 DR Pfam; PF02272; DHHA1; 1.  
 DR Pfam; PF01411; cRNA-synt\_2c; 1.  
 DR PRINTS; PR00980; TRNASYNTHALA.  
 DR TIGRFAMs; TIGR00344; alas; 1.  
 DR PROSITE; PSS0860; AA\_TRNA\_LIGASE\_II\_ALA; 1.  
 DR SEQUENCE 966 AA; 107741 MW; B1B63A7A8FB87B6E CRC64;  
 SQ  
 Query Match 37.7%; Score 52; DB 5; Length 966;  
 Best Local Similarity 50.0%; Pred. No. 19;  
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 QVTDHLEHYHLNVEMQWTT 19  
 Db 583 KVGDELEHLIDVERRWLT 600  
 RESULT 11  
 ID Q9U6B4 PRELIMINARY; PRT; 966 AA.  
 AC Q9U6B4;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Alanyl-cRNA synthetase.  
 GN AATS-ALA OR ALAS OR CG13391.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.

```

OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Chihade J.W., Brown J.R., Schimmel P., Ribas de Pouplana L.;
RT "Detection of an Intermediate Stage of Mitochondria Genesis.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF188718; AAF05593.1; -.
DR FlyBase; FBgn0027094; Aatc-ala.
DR GO; GO:0004813; P:alanine-tRNA ligase activity; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006419; P:alanyl-tRNA aminoacylation; IEA.
DR InterPro; IPR003156; DHHA1.
DR InterPro; IPR002318; tRNA-synt_2c.
DR InterPro; IPR006193; tRNA-synt_Ala.
DR Pfam; PF02272; DHHA1; 1.
DR Pfam; PF01411; tRNA-synt_2c; 1.
DR PRINTS; PR00980; TRNASYNTHALA.
DR TIGRPFAM; TIGR00344; ala5; 1.
DR PROSITE; PSS0860; AA_TRNA_LIGASE_II_ALA; 1.
KW Aminoacyl-tRNA synthetase.
SQ SEQUENCE 966 AA; 107877 MW; 82034A18F0C81E5 CRC64;

Query Match 37.7%; Score 52; DB 5; Length 966;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QVTDHLEYHLNVEMQWTT 19
Db 583 KVGDLEHLHIDVERHWT 600

RESULT 12
Q83XK8 PRELIMINARY; PRT; 678 AA.
ID Q83XK8;
AC Q83XK8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Leay protein.
GN LSSY.
OS Legionella pneumophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_Taxid=446;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Philadelphia;
RA Heuner K.;
RT "Description of a putative type I secretion system of Legionella
RT pneumophila.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ560770; CAD90962.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008934; AcPase_VanPerase.
DR InterPro; IPR000252; Deda.
DR InterPro; IPR000326; PA_PTPase.
DR Pfam; PF00597; Deda; 1.
DR Pfam; PF01569; PAP2; 1.
SQ SEQUENCE 678 AA; 76895 MW; AC817B2813F03DAF CRC64;

Query Match 37.3%; Score 51.5; DB 2; Length 678;
Best Local Similarity 44.0%; Pred. No. 16;
Matches 11; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

Qy 2 QVTDHLEYHLNVEMQWTTCKP 23
Db 480 QVHDHMPYHKEFTLTETKQWMDQRP 504

RESULT 13
Q8T1F5 PRELIMINARY; PRT; 2165 AA.
ID Q8T1F5

```

```

AC Q8T1F5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to Dictyostellium discoideum (Slime mold). protein tyrosine
DE kinase.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelid; Dictyostellium.
OX NCBI_Taxid=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szefranski K., Pachepat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungsai B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostellium discoideum.";
RL Nature 418:79-85 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116959; AAM08443.2; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001683; PK_Kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_kinase.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD0000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PSS0003; PH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PSS0195; PK; 1.
KW Kinase.
SQ SEQUENCE 2165 AA; 241910 MW; AFF8B5B1C32202F CRC64;

Query Match 37.3%; Score 51.5; DB 5; Length 2165;
Best Local Similarity 37.5%; Pred. No. 55;
Matches 9; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

Qy 2 QVTDHLEYHLNVEMQWTTCKP 22
Db 658 QIDDHIEYKQVLSNSNEWFCRR 681

RESULT 14
Q9YVN4 PRELIMINARY; PRT; 232 AA.
ID Q9YVN4;
AC Q9YVN4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF MSV208 putative uracil DNA glycosylase UNG (vaccinia DAR), similar
DE to Equine herpesvirus GB/U20824.
GN MSV208.
OS Melanoplus sanguinipes entomopoxvirus (MeEPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
OX NCBI_Taxid=83191;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tucson;

```

RX MEDLINE=99102612; PubMed=9847359;  
 RA Alfonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;  
 RT "The genome of Melanoplus sanguinipes entomopoxvirus";  
 RL J. Virol. 73:533-552(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tucson;  
 RA Alfonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;  
 RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF063866; AAC97753.1; -  
 DR PIR; T28369;  
 DR GO; GO:0004844; F:uracil DNA N-glycosylase activity; IEA.  
 DR GO; GO:0006281; P:DNA repair; IEA.  
 DR InterPro; IPR005122; UDNA\_glycylaseSF.  
 DR InterPro; IPR003249; U\_glycylase\_ncp.  
 DR Pfam; PF03167; UDG; 1.  
 DR PRODOM; PD001589; U\_glycylase\_ncp; 1.  
 SQ SEQUENCE 232 AA; 27297 MW; 85A9B79E1C4F09A0 CRC64;

Query Match 36.2%; Score 50; DB 12; Length 232;  
 Best Local Similarity 31.6%; Pred. No. 8.8;  
 Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Oy 2 QVTDHLEVHNVEMQWTTTC 20  
 Db 212 KANEYKIKHKNIEIDWISC 230

RESULT 15  
 O97N97  
 ID Q97N97 PRELIMINARY; PRT; 588 AA.  
 AC Q97N97;  
 DT 01-OCT-2001 (TREMblrel. 18, Created)  
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
 DE L-fucose isomerase.  
 GN SP2158.  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC BAA-334 / TIGR4;  
 RX MEDLINE=21357209; PubMed=11463916;  
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 RA Holtzapple E., Khouli H., Wolf A.M., Urdack T.R., Hansen C.L.,  
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,  
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
 RT "Complete genome sequence of a virulent isolate of Streptococcus  
 pneumoniae";  
 RT Science 293:498-506(2001).  
 RL EMBL; AE007504; AAK76212.1; -  
 DR PIR; C95252; C95252.  
 DR TIGR; SP2158; -  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0016853; F:isomerase activity; IEA.  
 DR GO; GO:0008736; F:L-fucose isomerase activity; IEA.  
 DR GO; GO:0006004; P:fucose metabolism; IEA.  
 DR InterPro; IPR004216; Fuc\_isomeraseC.  
 DR InterPro; IPR009015; Fuc\_isomeraseN.  
 DR Pfam; PF02952; fucose\_180\_C; 1.  
 DR isomerase; Complete proteome.  
 SQ SEQUENCE 588 AA; 65894 MW; 06E678EDF801786A CRC64;

Query Match 36.2%; Score 50; DB 16; Length 588;  
 Best Local Similarity 37.9%; Pred. No. 24;  
 Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Oy 1 RQVTDH-----LEVHNVEMQWTTCCXP 23  
 Db 301 RQVTDHFPNGDFMETFLNTQFDWNGIRKP 329

Search completed: March 18, 2004, 14:21:30  
 Job time : 16.381 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 29.9226 Seconds  
(without alignments)  
311.606 Million cell updates/sec

Title: US-09-941-314-12

Perfect score: 189  
Sequence: 1 QVTDHLEHYLNVEMQWTCCKPRTNCVPERE 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:.\*  
2: geneseqp1990s:.\*  
3: geneseqp2000s:.\*  
4: geneseqp2001s:.\*  
5: geneseqp2002s:.\*  
6: geneseqp2003as:.\*  
7: geneseqp2003bs:.\*  
8: geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189	100.0	33	AAU79862	AAU79862 Human cys
2	189	100.0	52	AAU79864	AAU79864 Human cys
3	189	100.0	59	AAU79866	AAU79866 Human cys
4	189	100.0	80	AAU79865	AAU79865 Human cys
5	189	100.0	115	AAU79853	AAU79853 Human cys
6	189	100.0	117	AAU79854	AAU79854 Human cys
7	189	100.0	137	AAU79852	AAU79852 Human cys
8	149	78.8	46	AAU79860	AAU79860 Human cys
9	149	78.8	49	AAU79863	AAU79863 Human cys
10	133	70.4	24	AAU79861	AAU79861 Human cys
11	133	70.4	48	AAU79867	AAU79867 Human cys
12	95	50.3	142	ADDA6708	ADDA6708 Rat Prote
13	95	50.3	142	ADDA6704	ADDA6704 Rat Prote
14	89	47.1	142	AAE02404	AAE02404 Murine cy
15	89	47.1	142	AAE04433	AAE04433 Mouse spe
16	89	47.1	143	ADA14374	ADA14374 Mouse spe
17	72.5	38.4	203	AAU83156	AAU83156 Novel sec
18	70	37.0	92	AAW78259	AAW78259 Fragment
19	70	37.0	123	AAW78260	AAW78260 Fragment
20	70	37.0	142	AAW78258	AAW78258 Fragment
21	70	37.0	142	AAE02405	AAE02405 Fragment
22	70	37.0	142	AAE04434	AAE04434 Human cys
23	70	37.0	142	ADA57231	ADA57231 Human sec
24	70	37.0	142	ADA41112	ADA41112 Human sec
25	70	37.0	142	ADC74335	ADC74335 Human sec

26	70	37.0	142	7	ADD37980	ADD37980 Human sec
27	70	37.0	142	7	ADD46706	ADD46706 Human Pro
28	70	37.0	142	7	ADD46710	ADD46710 Human Pro
29	68.5	36.2	138	4	AAE65889	AAE65889 Amino aci
30	67.5	35.7	37	4	ABBA4210	ABBA4210 Peptide #
31	67.5	35.7	37	4	AAAM35913	AAAM35913 Peptide #
32	67.5	35.7	37	4	ABBS25688	ABBS25688 Protein #
33	67.5	35.7	37	4	AAW75805	AAW75805 Human bon
34	67.5	35.7	37	4	AAAM62992	AAAM62992 Human bra
35	67.5	35.7	37	4	ABG57540	ABG57540 Human liv
36	67.5	35.7	37	5	ABGA5267	ABGA5267 Human pep
37	67	35.4	141	4	AAAY96576	AAAY96576 Murine cy
38	67	35.4	141	4	AAE02403	AAE02403 Murine cy
39	67	35.4	141	4	AAE04432	AAE04432 Mouse tes
40	65	34.4	113	6	ADA57563	ADA57563 Human sec
41	65	34.4	113	6	ADA41457	ADA41457 Human sec
42	65	34.4	113	7	ADC74577	ADC74577 Human sec
43	65	34.4	113	7	ADD38088	ADD38088 Human sec
44	65	34.4	114	2	AAW78153	AAW78153 Human sec
45	65	34.4	127	7	ADBE1282	ADBE1282 Rat Prote

## ALIGNMENTS

RESULT 1  
AAU79862  
ID AAU79862 standard; peptide; 33 AA.

XX AAU79862;

DT 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcy8) antigenic fragment #10.

XX Cystatin-8; Zcy8; cancer; procoagulant protein; thrombosis;

KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KW sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX W0200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026668.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting

XX spermatogenesis, and inhibiting cancer procoagulant protein which leads

XX to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 97; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcy8)

XX polypeptide (II). (I) is useful for: inhibiting cancer procoagulant

XX protein in an individual and thus inhibiting the thrombotic events

XX associated with cancer; promoting spermatogenesis, modulating seminal

XX fluid viscosity, enhancing viability of cryopreserved sperm, sperm

XX motility and fertilisation; and as antigenic peptides to generate

XX antibodies. Zcy8 is useful as research reagent for characterizing sites

XX of interaction between Zcy8 and its receptor. Zcy8 is useful in

XX enhancing fertilisation during assisted reproduction in humans and in

XX animals. Anti-(I) antibodies are useful to screen biological samples like

XX blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

XX presence of Zcy8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (1) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (1) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 33 AA;

Query Match 100.0%; Score 189; DB 5; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-18;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHLNVEMQWTTCKPPTNCVPOERE 33  
 Db 1 QVTDHLEHLNVEMQWTTCKPPTNCVPOERE 33

## RESULT 2

AAU79864  
 ID AAU79864 standard; peptide; 52 AA.

XX AAU79864;  
 XX  
 DT 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcys8) antigenic fragment #12.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 98; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(1) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (1) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (1) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 52 AA;

Query Match 100.0%; Score 189; DB 5; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-18;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHLNVEMQWTTCKPPTNCVPOERE 33  
 Db 20 QVTDHLEHLNVEMQWTTCKPPTNCVPOERE 52

## RESULT 3

AAU79866  
 ID AAU79866 standard; peptide; 59 AA.

XX AAU79866;

XX 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcys8) antigenic fragment #14.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 99; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(1) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (1) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (1) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 SQ Sequence 59 AA;

Query Match 100.0%; Score 189; DB 5; Length 59;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-18;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33  
 |||||  
 Db 2 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 34

RESULT 4  
 ID AAV79865 standard; peptide; 80 AA.  
 XX AAV79865;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #13.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.

OS Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026686.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 98; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)

CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal

CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate

CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in

CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8

CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene

CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX Sequence 80 AA;

Query Match 100.0%; Score 189; DB 5; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-18;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33  
 |||||  
 Db 23 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 55

RESULT 5  
 ID AAV79853 standard; protein; 115 AA.  
 XX AAV79853;

XX AAV79853;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #1.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic fragment.

OS Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026686.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 94; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)

CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal

CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate

CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in

CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8

CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene

CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)

XX Sequence 115 AA;

Query Match 100.0%; Score 189; DB 5; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-18;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 33  
 ID 53 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 85

RESULT 6  
 AAU79854  
 ID AAU79854 standard; protein; 117 AA.  
 AC AAU79854;  
 DT 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcys8) antigenic fragment #2.  
 DE  
 XX

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic fragment.

OS Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 94-95; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 CC  
 XX  
 SQ Sequence 117 AA;

Query Match 100.0%; Score 189; DB 5; Length 117;

Best Local Similarity 100.0%; Pred. No. 4,7e-18;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 33  
 ID 53 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 85

DB 55 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 87

RESULT 7  
 AAU79852  
 ID AAU79852 standard; protein; 137 AA.  
 AC AAU79852;  
 DT 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcys8).  
 DE  
 XX

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation.

OS Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

DR N-PSDB; ABK49522.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 93-94; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 137 AA;

Query Match 100.0%; Score 189; DB 5; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-18;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 33  
 ID 75 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 107

RESULT 8  
 ID AAU79860 standard; peptide: 46 AA.  
 XX  
 AC AAU79860;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #8.  
 XX  
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 97; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis; modulating seminal  
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 46 AA;  
 XX

Query Match 78.8%; Score 149; DB 5; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 5,2e-13;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEFHLNVEMQWTTCKPRTTN 26  
 |||||  
 Db 21 QVTDHLEFHLNVEMQWTTCKPRTTN 46

RESULT 9  
 ID AAU79863 standard; peptide: 49 AA.  
 XX

AC AAU79863;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #11.  
 XX  
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 97-98; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis; modulating seminal  
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 49 AA;  
 XX

Query Match 78.8%; Score 149; DB 5; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 5,6e-13;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEFHLNVEMQWTTCKPRTTN 26  
 |||||  
 Db 24 QVTDHLEFHLNVEMQWTTCKPRTTN 49

RESULT 10  
 ID AAU79861 standard; peptide: 24 AA.  
 XX  
 AC AAU79861;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX

DE Human cystatin-8 (Zcys8) antigenic fragment #9.  
 XX  
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001, 2001MO-US026868.  
 XX  
 PR 01-SEP-2000, 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 97; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterizing sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 24 AA;  
 Query Match 70.4%; Score 133; DB 5; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 4e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QVTDHLEHYLNVEWMQWTTCKQKE 23  
 Db 2 QVTDHLEHYLNVEWMQWTTCKQKE 24  
 RESULT 11  
 AAU79867  
 ID AAU79867 standard; peptide; 48 AA.  
 AC  
 XX AAU79867;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #15.  
 XX  
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KM sperm motility; fertilisation; antigenic peptide.  
 XX  
 XX Homo sapiens.  
 XX  
 PN W0200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001, 2001MO-US026868.  
 XX  
 PR 01-SEP-2000, 2000US-0230230P.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 DR WPI; 2002-383044/41.  
 XX  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 99; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterizing sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 48 AA;  
 Query Match 70.4%; Score 133; DB 5; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 NVEMQWTTCKQKEPTTNCVPOERE 33  
 Db 1 NVEMQWTTCKQKEPTTNCVPOERE 23  
 RESULT 12  
 ADD46708  
 ID ADD46708 standard; protein; 142 AA.  
 AC  
 XX ADD46708;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Rat Protein AAC36317, SEQ ID NO 12393.  
 XX  
 KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KM chronic constriction injury; CCI, spared nerve injury; SNI; Chung.  
 XX  
 OS Rattus norvegicus.  
 XX  
 PN W02003016475-A2.

```

XX 27-FEB-2003.
PD
PF 14-AUG-2002; 2002MO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
XX GENBANK; AAC36317.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 142 AA;
XX
Query Match 50.3%; Score 95; DB 7; Length 142;
Beet Local Similarity 45.5%; Pred. No. 4.6e-05;
Matches 15; Conservative 10; Mismatches 6; Indels 2; Gaps 1;
OY 1 QVTDHLEHLNANVMQMTTCOKP--ETTYCPOE 31
Db |||:||||::: :|||:||||:
77 QITDMEYHDIVQISRSNCRPLANTTENCIFOK 109
RESULT 13
ADD46704
ID ADD46704 standard; protein; 142 AA.
XX
XX ADD46704;
XX
XX 29-JAN-2004 (first entry)
XX
XX Rat Protein AAC36317, SEQ ID NO 12389.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

```

```

XX      Rattus norvegicus.
XX OS
XX PN WO2003016475-A2.
XX PD
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002MO-US025765.
XX PF
XX 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX PI
XX WPI, 2003-268312/26.
XX DR
XX GENBANK; AAC36317.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page: 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 142 AA;
XX
XX Query Match 50.3%; Score 95; DB 7; Length 142;
XX Best Local Similarity 45.5%; Pred. No. 4.6e-05;
XX Matches 15; Conservative 10; Mismatches 6; Indels 2; Gaps 1;
XX
XX 1 QVTDHLEHYLNAVEMQMTTCORP--ETTNVCVPOE 31
XX |||:||||::: : |||: |||:|:
XX 77 QITDRMEYHIDVQISRSNCRKPLANTENCIQK 109
XX
XX RESULT 14
XX ID AAE02404 standard; protein; 142 AA.
XX AC AAE02404;
XX
XX 10-AUG-2001 (first entry)
XX

```





SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-617-302-3

Query Match 47.1%; Score 89; DB 3; Length 142;  
Best Local Similarity 41.2%; Pred. No. 1.4e-05;  
Matches 14; Conservative 11; Mismatches 7; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCKP--ETNVCVPOE 32  
Db 77 QITDRMEYQIDVOISRSNCKKPLNTENCIPQK 110

RESULT 3  
US-09-431-480-4  
Sequence 4, Application US/09431480  
Patent No. 6235708  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72  
CURRENT APPLICATION NUMBER: US/09/431,480  
CURRENT FILING DATE: 1999-11-01  
EARLIER APPLICATION NUMBER: 60/109,217  
EARLIER FILING DATE: 1998-11-20  
EARLIER APPLICATION NUMBER: 60/156,382  
EARLIER FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-431-480-4

Query Match 37.0%; Score 70; DB 3; Length 142;  
Best Local Similarity 48.5%; Pred. No. 0.0085;  
Matches 16; Conservative 7; Mismatches 8; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCKP--ETNVCVPOE 31  
Db 77 QVNTLLEYLVLEVNIAIARTCKKPLSTNEICAIQE 109

RESULT 4  
US-09-617-302-4  
Sequence 4, Application US/09617302  
Patent No. 6245529  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72 C1  
CURRENT APPLICATION NUMBER: US/09/617,302  
CURRENT FILING DATE: 2000-07-17  
PRIOR APPLICATION NUMBER: 09/431,480  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: 60/109,217  
PRIOR FILING DATE: 1998-11-20  
PRIOR APPLICATION NUMBER: 60/156,382  
PRIOR FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-617-302-4

Query Match 37.0%; Score 70; DB 3; Length 142;  
Best Local Similarity 48.5%; Pred. No. 0.0085;  
Matches 16; Conservative 7; Mismatches 8; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCKP--ETNVCVPOE 31  
Db 77 QVNTLLEYLVLEVNIAIARTCKKPLSTNEICAIQE 109

RESULT 5  
US-09-431-480-2  
Sequence 2, Application US/09431480  
Patent No. 6235708  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72  
CURRENT APPLICATION NUMBER: US/09/431,480  
CURRENT FILING DATE: 1999-11-01  
EARLIER APPLICATION NUMBER: 60/109,217  
EARLIER FILING DATE: 1998-11-20  
EARLIER APPLICATION NUMBER: 60/156,382  
EARLIER FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 141  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-431-480-2

Query Match 35.4%; Score 67; DB 3; Length 141;  
Best Local Similarity 42.4%; Pred. No. 0.023;  
Matches 14; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCKP--ETNVCVPOE 31  
Db 76 QITDSLEYLVLEVNIAIARTCKKIKAGDNENCLFQ 108

RESULT 6  
US-09-617-302-2  
Sequence 2, Application US/09617302  
Patent No. 6245529  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T  
FILE REFERENCE: 98-72 C1  
CURRENT APPLICATION NUMBER: US/09/617,302  
CURRENT FILING DATE: 2000-07-17  
PRIOR APPLICATION NUMBER: 09/431,480  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: 60/109,217  
PRIOR FILING DATE: 1998-11-20  
PRIOR APPLICATION NUMBER: 60/156,382  
PRIOR FILING DATE: 1999-09-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 141  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-617-302-2

Query Match 35.4%; Score 67; DB 3; Length 141;  
Best Local Similarity 42.4%; Pred. No. 0.023;  
Matches 14; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCKP--ETNVCVPOE 31  
Db 76 QITDSLEYLVLEVNIAIARTCKKIKAGDNENCLFQ 108

RESULT 7  
US-08-849-303-19  
Sequence 19, Application US/08849303  
Patent No. 6680424  
GENERAL INFORMATION:  
APPLICANT: Atkinson, Howard J.  
APPLICANT: McPherson, Michael J.  
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/849,303  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1321-1-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 127 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-849-303-19

Query Match 34.4%; Score 65; DB 4; Length 127;  
Best Local Similarity 48.3%; Pred. No. 0.04;  
Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYHLNVEQMWTTCOKPET--TNC 27  
Db 62 QLVAGINYYFLDVEWGRITTCRSQTNLTNC 90

RESULT 8  
US-09-886-319A-46  
Sequence 46, Application US/09886319A  
Patent No. 6586185  
GENERAL INFORMATION:  
APPLICANT: Wolf, Eckard  
APPLICANT: Werner, Sabine  
APPLICANT: Halle, Jörn-Peter  
APPLICANT: Regenbogen, Johannes  
TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
the Diagnosis or Treatment of Skin Disorders and Wound  
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically  
TITLE OF INVENTION: Active Substances  
FILE REFERENCE: 50125/014002  
CURRENT APPLICATION NUMBER: US/09/886,319A  
CURRENT FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: US 60/222,081

PRIOR FILING DATE: 2000-08-01  
PRIOR APPLICATION NUMBER: DE 10030149.5  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 46  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-886-319A-46

Query Match 29.6%; Score 56; DB 4; Length 140;  
Best Local Similarity 44.8%; Pred. No. 0.95;  
Matches 13; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYHLNVEQMWTTCOKPET--TNC 27  
Db 75 QLVAGVNYFLDVEWGRITTCRSQTNLTDC 103

RESULT 9  
US-09-886-319A-48  
Sequence 48, Application US/09886319A  
Patent No. 6586185  
GENERAL INFORMATION:  
APPLICANT: Wolf, Eckard  
APPLICANT: Werner, Sabine  
APPLICANT: Halle, Jörn-Peter  
APPLICANT: Regenbogen, Johannes  
TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for  
the Diagnosis or Treatment of Skin Disorders and Wound  
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically  
TITLE OF INVENTION: Active Substances  
FILE REFERENCE: 50125/014002  
CURRENT APPLICATION NUMBER: US/09/886,319A  
CURRENT FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: US 60/222,081  
PRIOR FILING DATE: 2000-08-01  
PRIOR APPLICATION NUMBER: DE 10030149.5  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 48  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-886-319A-48

Query Match 29.6%; Score 56; DB 4; Length 140;  
Best Local Similarity 44.8%; Pred. No. 0.95;  
Matches 13; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYHLNVEQMWTTCOKPET--TNC 27  
Db 75 QLVAGVNYFLDVEWGRITTCRSQTNLTDC 103

RESULT 10  
US-08-849-303-26  
Sequence 26, Application US/08849303  
Patent No. 6680424  
GENERAL INFORMATION:  
APPLICANT: Atkinson, Howard J.  
APPLICANT: McPherson, Michael J.  
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey

COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/849,303  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1321-1-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-849-303-26

Query Match 28.6%; Score 54.5; DB 4; Length 111;  
Best Local Similarity 33.3%; Pred. No. 1.2;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC-VPOERE 33  
DB 47 QVAVGVVFLDVELGRTTKTQPNLNC 88

RESULT 11  
US-09-775-932-2  
Sequence 2, Application US/09775932  
Patent No. 6534477  
GENERAL INFORMATION:  
APPLICANT: University of British Columbia  
TITLE OF INVENTION: Production and use of Modified Cycatins  
FILE REFERENCE: 58069  
CURRENT APPLICATION NUMBER: US/09/775,932  
CURRENT FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: CA99/00717  
PRIOR FILING DATE: 1999-08-05  
PRIOR APPLICATION NUMBER: 60/095,503  
PRIOR FILING DATE: 1998-08-05  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 120  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-775-932-2

Query Match 28.6%; Score 54; DB 4; Length 120;  
Best Local Similarity 41.4%; Pred. No. 1.6;  
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC 27  
DB 55 QVAVGVVFLDVELGRTTKTQPNLNC 83

RESULT 12  
5432264-4  
Patent No. 5432264  
APPLICANT: GRUBB, ANDERS; LUNDWALL, AKE; ABRAHAMSON, MAGNUS;

DALBOGE, HENRIK  
TITLE OF INVENTION: RECOMBINANT 3-DES-OH-CYSTATIN C PRODUCED  
BY EXPRESSION IN A PROCAROTIC HOST CELL  
NUMBER OF SEQUENCES: 8  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/929,290  
FILING DATE: 13-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 440,221  
FILING DATE: 21-NOV-1989  
APPLICATION NUMBER: 297,198  
FILING DATE: 20-MAY-1988  
SEQ ID NO: 4  
LENGTH: 120  
5432264-4

Query Match 28.6%; Score 54; DB 6; Length 120;  
Best Local Similarity 41.4%; Pred. No. 1.6;  
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC 27  
DB 55 QVAVGVVFLDVELGRTTKTQPNLNC 83

RESULT 13  
US-08-832-535-11  
Sequence 11, Application US/08832535  
Patent No. 5919658  
GENERAL INFORMATION:  
APPLICANT: NI, JIAN  
APPLICANT: LI, HAODONG  
APPLICANT: YU, GUO-LIANG  
APPLICANT: GENTZ, REINER L  
TITLE OF INVENTION: HUMAN CYSTATIN F  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
STREET: 9410 KEY WEST AVENUE  
CITY: ROCKVILLE  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/832,535  
FILING DATE: 03-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KIMBALL, PAUL C.  
REGISTRATION NUMBER: 34,610  
REFERENCE/DOCKET NUMBER: PF265  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 994-1700  
TELEFAX: (201) 994-1744  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 145 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-832-535-11

Query Match 28.6%; Score 54; DB 2; Length 145;  
Best Local Similarity 41.4%; Pred. No. 1.9;  
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC 27

Db 81 QIVAGVNYFLDVELGRTTCTKTPNLDNC 109

RESULT 14  
US-08-791-522-3  
Sequence 3, Application US/08791522  
Patent No. 5935817  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goll, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE  
PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/791.522  
FILING DATE: Filed Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0193 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 146 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 181387  
US-08-791-522-3

Query Match 28.6%; Score 54; DB 2; Length 146;  
Best Local Similarity 41.4%; Pred. No. 2;  
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETNC 27  
Db 81 QIVAGVNYFLDVELGRTTCTKTPNLDNC 109

RESULT 15  
US-08-744-138-3  
Sequence 3, Application US/08744138  
Patent No. 6011012  
GENERAL INFORMATION:  
APPLICANT: Gentz, Retner L.  
APPLICANT: Ni, Jian  
APPLICANT: Rosen, Craig A.  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: Human Cystatin E  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/744.138  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF202P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301 309 8504  
TELEFAX: 301 309 8512  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 146 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Cystatin C  
US-08-744-138-3

Query Match 28.6%; Score 54; DB 3; Length 146;  
Best Local Similarity 41.4%; Pred. No. 2;  
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETNC 27  
Db 81 QIVAGVNYFLDVELGRTTCTKTPNLDNC 109

Search completed: March 18, 2004, 14:25:37  
Job time : 9.44643 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 19.9702 Seconds

(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-12

Perfect score: 189

Sequence: 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubppaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubppaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	33	9	US-09-941-314-12 Sequence 12, Appl
2	189	100.0	52	9	US-09-941-314-14 Sequence 14, Appl
3	189	100.0	59	9	US-09-941-314-16 Sequence 15, Appl
4	189	100.0	80	9	US-09-941-314-15 Sequence 16, Appl
5	189	100.0	115	9	US-09-941-314-3 Sequence 3, Appl
6	189	100.0	117	9	US-09-941-314-4 Sequence 4, Appl
7	189	100.0	137	9	US-09-941-314-2 Sequence 2, Appl
8	149	78.8	46	9	US-09-941-314-10 Sequence 10, Appl
9	149	78.8	49	9	US-09-941-314-13 Sequence 13, Appl
10	133	70.4	24	9	US-09-941-314-11 Sequence 11, Appl
11	133	70.4	48	9	US-09-941-314-17 Sequence 17, Appl
12	72.5	38.4	203	9	US-09-893-737-164 Sequence 164, App
13	68.5	36.2	138	14	US-10-239-663-37 Sequence 37, Appl
14	67.5	35.7	37	9	US-09-864-761-40986 Sequence 40986, A
15	65	34.4	127	8	US-08-849-303-19 Sequence 19, Appl

16	63	33.3	145	9	US-09-740-638-2	Sequence 2, Appl
17	63	33.3	145	13	US-10-006-467-2	Sequence 2, Appl
18	63	33.3	145	14	US-10-235-148-2	Sequence 2, Appl
19	63	33.3	145	14	US-10-168-425-14	Sequence 14, Appl
20	63	33.3	165	9	US-09-740-638-5	Sequence 5, Appl
21	63	33.3	165	13	US-10-006-467-5	Sequence 5, Appl
22	63	33.3	165	13	US-10-235-148-5	Sequence 5, Appl
23	56	29.6	140	14	US-10-376-564-46	Sequence 46, Appl
24	56	29.6	140	14	US-10-376-564-48	Sequence 48, Appl
25	55.5	29.4	633	10	US-09-911-458-2	Sequence 2, Appl
26	55.5	29.4	633	14	US-10-298-974-2	Sequence 2, Appl
27	54.5	28.8	111	8	US-08-849-303-26	Sequence 26, Appl
28	54	28.6	120	9	US-09-775-932-2	Sequence 2, Appl
29	54	28.6	146	8	US-08-849-303-17	Sequence 17, Appl
30	54	28.6	146	9	US-09-940-497-4	Sequence 3, Appl
31	54	28.6	146	9	US-09-969-834-3	Sequence 3, Appl
32	54	28.6	146	14	US-10-329-428-3	Sequence 3, Appl
33	54	28.6	146	14	US-10-376-564-47	Sequence 47, Appl
34	53	28.0	122	9	US-09-775-932-10	Sequence 10, Appl
35	53	28.0	142	8	US-08-849-303-20	Sequence 20, Appl
36	53	28.0	142	9	US-09-940-497-4	Sequence 4, Appl
37	53	28.0	142	12	US-10-262-839-132	Sequence 132, App
38	53	28.0	604	9	US-09-820-809-13	Sequence 13, Appl
39	53	28.0	983	14	US-10-255-149-2	Sequence 2, Appl
40	52	27.5	140	8	US-08-849-303-18	Sequence 18, Appl
41	51.5	27.2	678	12	US-10-282-1228-61129	Sequence 61129, A
42	51	27.0	447	12	US-10-424-599-191984	Sequence 191984, A
43	50	26.5	109	14	US-10-029-386-30645	Sequence 16, Appl
44	50	26.5	112	8	US-08-849-303-16	Sequence 24, Appl
45	50	26.5	118	9	US-09-775-932-24	

## ALIGNMENTS

RESULT 1

US-09-941-314-12

Sequence 12, Application US/09941314

Patent No. US20020142396A1

GENERAL INFORMATION:

APPLICANT: ZymoGenetics, Inc.

TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to Inhibit Cancer Procoagulant Protein

FILE REFERENCE: 00-81PC

CURRENT APPLICATION NUMBER: US/09/941,314

CURRENT FILING DATE: 2001-08-29

PRIOR APPLICATION NUMBER: 60/230,230

PRIOR FILING DATE: 2001-09-01

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 12

LENGTH: 33

TYPE: PRT

ORGANISM: Homo sapiens

US-09-941-314-12

Query Match 100.0%; Score 189; DB 9; Length 33;

Best Local Similarity 100.0%; Pred. No. 1e-18;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33

1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33

RESULT 2

US-09-941-314-14

Sequence 14, Application US/09941314

Patent No. US20020142396A1

GENERAL INFORMATION:

APPLICANT: ZymoGenetics, Inc.

TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to Inhibit Cancer Procoagulant Protein

FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 52  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-14

Query Match 100.0%; Score 189; DB 9; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1,7e-18;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33  
Db 20 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 52

RESULT 3  
US-09-941-314-16  
Sequence 16, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16  
LENGTH: 59  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-16

Query Match 100.0%; Score 189; DB 9; Length 59;  
Best Local Similarity 100.0%; Pred. No. 1,9e-18;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33  
Db 2 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 34

RESULT 4  
US-09-941-314-15  
Sequence 15, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-15

Query Match 100.0%; Score 189; DB 9; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2,7e-18;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33  
Db 23 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 55

RESULT 5  
US-09-941-314-3  
Sequence 3, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 115  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-3

Query Match 100.0%; Score 189; DB 9; Length 115;  
Best Local Similarity 100.0%; Pred. No. 4e-18;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33  
Db 53 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 85

RESULT 6  
US-09-941-314-4  
Sequence 4, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 117  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-4

Query Match 100.0%; Score 189; DB 9; Length 117;  
Best Local Similarity 100.0%; Pred. No. 4,1e-18;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33  
Db 55 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 87

RESULT 7  
US-09-941-314-2  
Sequence 2, Application US/09941314  
Patent No. US20020142396A1

```
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-2

Query Match          100.0%; Score 189; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 4,8e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVTDHLEHYHLNVEMQWTTCKPPTNCVPOERE 33
Db 75 QVTDHLEHYHLNVEMQWTTCKPPTNCVPOERE 107

RESULT 8
US-09-941-314-10
; Sequence 10, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-10

Query Match          78.8%; Score 149; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 4,4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVTDHLEHYHLNVEMQWTTCKPPTN 26
Db 21 QVTDHLEHYHLNVEMQWTTCKPPTN 46

RESULT 9
US-09-941-314-13
; Sequence 13, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 49
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-13

Query Match          78.8%; Score 149; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 4,7e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVTDHLEHYHLNVEMQWTTCKPPTN 26
Db 24 QVTDHLEHYHLNVEMQWTTCKPPTN 49

RESULT 10
US-09-941-314-11
; Sequence 11, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-11

Query Match          70.4%; Score 133; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 3,4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVTDHLEHYHLNVEMQWTTCKP 23
Db 2 QVTDHLEHYHLNVEMQWTTCKP 24

RESULT 11
US-09-941-314-17
; Sequence 17, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-17

Query Match          70.4%; Score 133; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 7,1e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 NVEMQWTTCKPPTNCVPOERE 33
Db 1 NVEMQWTTCKPPTNCVPOERE 23
```

RESULT 12  
US-09-893-737-164  
; Sequence 164, Application US/09893737  
; Patent No. US20020110855A1  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Presnell, Scott R.  
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS  
; FILE REFERENCE: 00-41  
; CURRENT APPLICATION NUMBER: US/09/893,737  
; PRIOR FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: US 60/215,446  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 329  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 164  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-893-737-164

Query Match 38.4%; Score 72.5; DB 9; length 203;  
Best Local Similarity 50.0%; Pred. No. 0.064; Indels 1; Gaps 1;  
Matches 16; Conservative 5; Mismatches 10;

Qy 1 QVTDLEHYHNVEMQWTTQCK-PEITNCVPOE 31  
Db 115 QVTDLEHYIEVKIARTICKKISENCAFOE 146

RESULT 13  
US-10-239-663-37  
; Sequence 37, Application US/10239663  
; Publication No. US20030139572A1  
; GENERAL INFORMATION:  
; APPLICANT: Agarwal, Pankaj  
; APPLICANT: Murdoch, Paul R.  
; APPLICANT: Rizvi, Safia, K.  
; APPLICANT: Smith, Randall, F.  
; APPLICANT: Xiang, Zhaoying  
; APPLICANT: Kahnick, Karen  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; FILE REFERENCE: GPO50018  
; CURRENT APPLICATION NUMBER: US/10/239,663  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: PCT/US01/09226  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: 60/192,158  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,668  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: 60/200,166  
; PRIOR FILING DATE: 2000-04-27  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 138  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-239-663-37

Query Match 36.2%; Score 68.5; DB 14; length 138;  
Best Local Similarity 46.9%; Pred. No. 0.15; Indels 1; Gaps 1;  
Matches 15; Conservative 6; Mismatches 10;

Qy 1 QVTDLEHYHNVEMQWTTQCK-PEITNCVPOE 31  
Db 74 KVTDSEYIEYIEVKIARTICKKISENCAFOE 105

RESULT 14  
US-09-864-761-40986  
; Sequence 40986, Application US/09864761

; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wenheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263,6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 40986  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL121894.14  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
; OTHER INFORMATION: EST HUMAN HIT: AW665128.1, EVALUATE 1.10e-02  
; OTHER INFORMATION: SWISSPROT HIT: O88969, EVALUATE 1.00e-04  
US-09-864-761-40986

Query Match 35.7%; Score 67.5; DB 9; length 37;  
Best Local Similarity 48.4%; Pred. No. 0.05; Indels 1; Gaps 1;  
Matches 15; Conservative 5; Mismatches 10;

Qy 2 VTDLLEHYHNVEMQWTTQCK-PEITNCVPOE 31  
Db 1 VTDSLEHYIEVKIARTICKKISENCAFOE 31





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15, Search time 6.35119 Seconds  
(without alignments)  
499.799 Million cell updates/sec

Title: US-09-941-314-12

Perfect score: 189  
Sequence: 1 QVTDHLEHYHNVEMQWTTCKPRTNCPQERE 33

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	47.1	139	2 A45361	cytostatin-related e
2	65	34.4	120	2 S10587	cytostatin C - rat
3	65	34.4	127	2 S07085	cytostatin C precurs
4	64	33.9	1779	2 T23130	hypothetical prote
5	55	29.1	4540	2 T30838	cytoplasmic dynein
6	54.5	28.8	111	2 A28793	cytostatin - putf ad
7	54	28.6	146	1 UDHU	cytostatin C precurs
8	54	28.6	302	2 B96520	hypothetical prote
9	54	28.6	463	2 E86464	hypothetical prote
10	53	28.0	142	2 A47142	cytostatin D precurs
11	53	28.0	912	2 JH0563	metabotropic gluta
12	53	28.0	3766	2 T29185	hypothetical prote
13	52	27.5	140	2 A36163	cytostatin C precurs
14	51.5	27.2	371	2 B96614	hypothetical prote
15	50	26.5	112	1 UDBO	cytostatin - bovine
16	50	26.5	232	2 T28369	probable uracil-DN
17	50	26.5	717	1 XNBVCM	glutamine-fructose
18	50	26.5	725	2 T42688	hypothetical prote
19	49.5	26.2	622	2 S63539	GABA/beta-alanine
20	49.5	26.2	953	2 T12577	H-exporting ATPase
21	49.5	26.2	1073	2 F89467	protein R09H3.1 [i
22	49	25.9	133	2 B86754	prophage p12 prote
23	49	25.9	328	2 S01359	salivary glue prot
24	49	25.9	457	2 C86454	hypothetical prote
25	49	25.9	695	2 D84634	hypothetical prote
26	49	25.9	1905	2 I51553	Plexin - African c
27	48.5	25.7	331	2 T41758	ACMNPV orf13 - Bom
28	48	25.4	133	2 JC4536	cytostatin precursor
29	48	25.4	455	2 T00856	pectate lyase (EC

30	48	25.4	603	2 T33134	hypothetical prote
31	47.5	25.1	214	1 A8BYD	ADP8 protein - Yea
32	47.5	25.1	1100	2 T47940	hypothetical prote
33	47.5	25.1	1275	2 T49362	hypothetical prote
34	47.5	25.1	1308	2 B75198	DNA helicase relat
35	47	24.9	180	2 G69222	molypdenum formylm
36	47	24.9	447	2 T19078	hypothetical prote
37	47	24.9	456	2 E84960	adenylosuccinate 1
38	47	24.9	555	2 S39953	isocitrate lyase (
39	47	24.9	610	2 T47725	hypothetical prote
40	47	24.9	1554	2 T06370	probable DNA (cyto
41	47	24.9	3161	2 T30342	protein HMWP1 - Ye
42	46.5	24.6	239	2 AB1834	pyridoxal phosphat
43	46.5	24.6	713	2 JC6012	glutamine-fructose
44	46.5	24.6	764	2 A84456	mutator-like trans
45	46	24.3	106	2 T27989	hypothetical prote

## ALIGNMENTS

### RESULT 1

A45361  
cytostatin-related epididymal specific protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999

C:Accession: A45361

R:Corwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.

Mol. Endocrinol. 6, 1653-1664, 1992

A>Title: The CRE5 gene: a unique testis-regulated gene related to the cytostatin family is

A:Reference number: A45361; MUID:3078799; PMID:1280328

A:Accession: A45361

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Cross-references: GB:849926; NID:9260492; PIDN:AAC35390.1; PID:9260493

A>Note: sequence extracted from NCBI backbone (NCBIP:118813)

C:Superfamily: cytostatin; cytostatin homology

F:28-139/Domain: cytostatin homology <CTS>

Query Match 47.1%; Score 89; DB 2; Length 139;  
Best Local Similarity 41.2%; Pred. No. 6.1e-05;  
Matches 14; Conservative 11; Mismatches 7; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYHNVEMQWTTCKP--ETNVCVPOER 32

Db 74 QVTDHLEHYHNVEMQWTTCKP--ETNVCVPOER 107

### RESULT 2

S10587  
cytostatin C - rat

C:Species: Rattus sp. (rat)

C>Date: 21-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 16-Jul-1999

C:Accession: S10587

R:Bernard, F.; Bernard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Billard, M.; Gauch

Biol. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990

A>Title: Rat cytostatin C: the complete amino acid sequence reveals a site for N-glycosyla

A:Reference number: S10587; MUID:90380276; PMID:2400577

A:Accession: S10587

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-120 <ESN>

A>Note: 43-Asn was also found

C:Superfamily: cytostatin; cytostatin homology

F:9-120/Domain: cytostatin homology <CVS>

Query Match 34.4%; Score 65; DB 2; Length 120;  
Best Local Similarity 48.3%; Pred. No. 0.098;  
Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYHNVEMQWTTCKP--TNC 27

Db 55 QLVAGINYYLDVEMGRITTCSTQNTLINC 83

## RESULT 3

S07085

cystatin C precursor - rat (fragment)

C/Species: Rattus norvegicus (Norway rat)

C/Date: 01-Dec-1993 #sequence\_revision 03-Aug-1995 #text\_change 16-Jul-1999

C/Accession: S07085; S01337; S21109

R/Cole, T.; Dickson, P.W.; Esnard, F.; Averill, S.; Risbridger, G.P.; Gauthier, F.; Sch

Eur. J. Biochem. 186, 35-42, 1989

A/Title: The cDNA structure and expression analysis of the genes for the cysteine prote

A/Reference number: S07085; MUID:90092122; PMID:2689174

A/Accession: S07085

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-127 &lt;COL&gt;

A/Cross-references: EMBL:X16957; NID:956041; PIDN:CAA34831.1; PID:G736290

R/Esnard, A.; Esnard, F.; Faucher, D.; Gauthier, F.

FEBS Lett. 236, 475-478, 1988

A/Title: Two rat homologues of human cystatin C.

A/Reference number: S01337; MUID:88313020; PMID:3044831

A/Accession: S01337

A/Molecule type: protein

A/Residues: 8-49 &lt;ESN&gt;

R/Esnard, A.; Esnard, F.; Guillou, F.; Gauthier, F.

FEBS Lett. 300, 131-135, 1992

A/Title: Production of the cysteine proteinase inhibitor cystatin C by rat Sertoli cells

A/Reference number: S21109; MUID:9225121; PMID:1563513

A/Accession: S21109

A/Molecule type: protein

A/Residues: 8, 'XX', 11-20 &lt;ES2&gt;

C/Superfamily: cystatin; cystatin homology

C/Keywords: cysteine proteinase inhibitor

F/16-127/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

F/80-90, 104-124/Dmain: cystatin homology &lt;CTS&gt;

A/Map position: 1  
A/Intons: 161/2; 223/2; 309/3; 332/2; 547/3; 603/1; 657/3; 745/2; 802/1; 856/1; 1031/1;  
C/Superfamily: Caenorhabditis elegans hypothetical protein T28B8.3

Query Match 33.9%; Score 64; DB 2; Length 1779;  
Best Local Similarity 34.2%; Pred. No. 2.1;  
Matches 13; Conservative 8; Mismatches 7; Indels 10; Gaps 2;

Db 750 RITDHDYIINNPLKHPNVQMIT--DPTTHCLPK 785

## RESULT 5

T30838

cytoplasmic dynein heavy chain - Parametium tetraurelia

C/Species: Parametium tetraurelia

C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 02-Mar-2001

C/Accession: T30838; PC4341

R/Kandl, K.A.; Forney, J.D.; Asai, D.J.

submitted to the EMBL Data Library, January 1995

A/Description: The dynein genes of Parametium: the differential expression of axonemal ar

A/Reference number: Z20502

A/Accession: T30838

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-4540 &lt;KAN&gt;

R/Asai, D.J.; Beckwith, S.M.; Kandl, K.A.; Keating, H.H.; Tjandra, H.; Forney, J.D.

J. Cell Sci. 107, 839-847, 1994

A/Title: The dynein genes of Parametium tetraurelia: Sequences adjacent to the catalytic

A/Reference number: PC4340; MUID:94334383; PMID:8056840

A/Accession: PC4341

A/Molecule type: mRNA

A/Residues: 1831-2029 &lt;ASA&gt;

C/Genetics: code: SGC5

A/Intons: 57/3

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

Query Match 28.8%; Score 54.5; DB 2; Length 111;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVRYIIMBELKTTCKTVGRPKGYEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVRYIIMBELKTTCKTVGRPKGYEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVRYIIMBELKTTCKTVGRPKGYEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVRYIIMBELKTTCKTVGRPKGYEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVRYIIMBELKTTCKTVGRPKGYEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVRYIIMBELKTTCKTVGRPKGYEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVRYIIMBELKTTCKTVGRPKGYEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVRYIIMBELKTTCKTVGRPKGYEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVRYIIMBELKTTCKTVGRPKGYEIQNCNLPENQ 88

## RESULT 7

UDHU

Cystatin C precursor [validated] - human

N/Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma C; Species: Homo sapiens (man)

C/Date: 06-Jul-1982 #sequence revision 31-Mar-1991 #text change 08-Dec-2000

C/Accession: S10216; S00004; J10095; A33400; S02751; A01270; A25434; S12288; A32732; A60 R; Abrahamson, M.; Olafsson, I.; Paldestedt, A.; Ulfvbaeck, M.; Lundwall, A.; Jansson, O.

Biochem. J. 268, 287-294, 1990

A/Title: Structure and expression of the human cystatin C gene.

A/Reference number: S10216; MUID:90303202; PMID:2363674

A/Accession: S10216

A/Molecule type: DNA

A/Residues: 1-146 &lt;AB1&gt;

A/Cross-references: EMBL:X52255; NID:g30357; PIND:CAA36497.1; PID:g296643

R; Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.

FEBS Lett. 216, 229-233, 1987

A/Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the

A/Reference number: S00004; MUID:67219149; PMID:3495457

A/Accession: S00004

A/Molecule type: mRNA

A/Residues: 1-146 &lt;AB2&gt;

A/Cross-references: EMBL:X05607; NID:g30371; PIND:CAA29096.1; PID:g755738

R; Levy, E.; Lopez-Otin, C.; Ghiso, J.; Gellner, D.; Frangione, B.

J. Exp. Med. 169, 1771-1778, 1989

A/Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a

A/Reference number: J10095; MUID:89235594; PMID:2541223

A/Accession: J10095

A/Molecule type: DNA

A/Residues: 1-146 &lt;LEV&gt;

A/Cross-references: GB:X61681; NID:g30367; PIND:CAA43856.2; PID:g4490944

A/Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit

e)

R; Saitoh, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada, Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989

A/Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which

A/Reference number: A33400; MUID:89350949; PMID:2764935

A/Accession: A33400

A/Molecule type: DNA

A/Residues: 1-24, 'T', 26-146 &lt;SA1&gt;

A/Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:g181385; PIND:AAA52164.1; PID:g

R; Ghiso, J.; Cowan, N.; Frangione, B.

Biol. Chem. Hoppe-Seyler 369, 205-208, 1988

A/Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron

A/Reference number: S02751; MUID:89076507; PMID:3264504

A/Accession: S02751

A/Molecule type: DNA

A/Residues: 82-119 &lt;GH2&gt;

A/Cross-references: EMBL:M27769

A/Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown

R; Grubb, A.; Lofberg, H.

Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982

A/Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in th

A/Reference number: A01270; MUID:82222268; PMID:6283552

A/Accession: A01270

A/Molecule type: protein

A/Residues: 27-131, 'S', 133-146 &lt;GRU&gt;

R; Ghiso, J.; Jansson, O.; Frangione, B.

Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986

A/Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland t

A/Reference number: A25434; MUID:66206076; PMID:3517880

A/Accession: A25434

A/Molecule type: protein

A/Residues: 37-93, 'Q', 95-146 &lt;GHI&gt;

R; Turk, V.; Brzin, J.; Longner, M.; Ritonja, A.; Brodskin, M.; Borchart, U.; Machleidt, W.

Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983

A/Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystati

A/Reference number: S01461; MUID:64110055; PMID:6662498

A/Accession: S12288

A/Molecule type: protein

A/Residues: 27-73 &lt;TUR&gt;

R; Brzin, J.; Popovic, T.; Turk, V.

Biochem. Biophys. Res. Commun. 118, 103-109, 1984

A/Title: Human cystatin, a new protein inhibitor of cysteine proteinases.

A/Reference number: A32732; MUID:84128015; PMID:6365094

A/Accession: A32732

A/Molecule type: protein

A/Residues: 27-76 &lt;BR2&gt;

R; Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.

Scand. J. Clin. Lab. Invest. 50, 85-93, 1990

A/Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst.

A/Reference number: A60552; MUID:90193615; PMID:2315647

A/Accession: A60552

A/Molecule type: protein

A/Residues: 27-49, 'XX', 52-64 &lt;OLA&gt;

A/Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d

e defective gene is not present in CSF but is found instead in amyloid deposits

R; Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.

Biol. Chem. Hoppe-Seyler 371, 575-580, 1990

A/Title: Different forms of human cystatin C.

A/Reference number: S10607; MUID:91025625; PMID:2222856

A/Accession: S10607

A/Molecule type: protein

A/Residues: 27-53 &lt;POP&gt;

A/Experimental source: urine, kidney disease

A/Note: truncated forms with amino ends at positions 35 and 36 of the precursor were also

R; Grubb, A.; Lofberg, H.; Barrett, A.J.

FEBS Lett. 170, 370-374, 1984

A/Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.

A/Reference number: S01462

A/Accession: S01462

A/Contents: annotation; disulfide bonds

R; Bertl, P.J.; Storer, A.C.

Biochem. J. 302, 411-416, 1994

A/Title: Local pH-dependent conformational changes leading to proteolytic susceptibility

A/Reference number: S55305; MUID:94379969; PMID:8092991

A/Accession: S55305

A/Status: preliminary

A/Molecule type: protein

A/Residues: 27-49/106-146 &lt;BER&gt;

C/Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl

e patients with certain autoimmune diseases.

C/Comment: This protein is an inhibitor of cysteine proteinases and may serve an import

C/Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorr

C/Genetics:

A/Gene: GDB:CST3

A/Cross-references: GDB:119817; OMIM:105150

A/Map position: 20p11.2-20p11.2

A/Intons: 81/3, 119/3

C/Superfamily: cystatin; cystatin homology

C/Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline

F:1-26/Domain: signal sequence #status predicted &lt;SIG&gt;

F:27-146/Product: cystatin C #status experimental &lt;MAT&gt;

F:35-146/Domain: cystatin homology &lt;CTS&gt;

F:81-85/Region: inhibitory #status predicted

F:29/Modified site: hydroxyproline (Pro) (partial) #status experimental

F:99-109,123-143/Disulfide bonds: #status experimental

Query Match	28.6%	Score 54;	DB 1;	Length 146;
Best Local Similarity	41.4%	Pred. No. 3.8;		
Matches 12;	Conservative 4;	Mismatches 11;	Indels 2;	Gaps 1;

Qy	1 QVTDHLEHYHNVEMQWTTQCK--PETTNC 27
Db	81 QIVAGVYVFLDVELGRTTKTKTQPNLNC 109

## RESULT 8

B96520

Hydrochemical protein T2915.11 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C/Accession: B96520

R; Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hultzar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Rowley, D.; Sakano, H.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: B96520  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-302 <STO>  
 A/Cross-references: GB:AE005173; NID:g10645473; PIDN:AA621587.1; GSPDB:GN00141  
 A/Genetics: 1  
 A/Map position: 1

Query Match 28.6%; Score 54; DB 2; Length 302;  
 Best Local Similarity 42.3%; Pred. No. 8;  
 Matches 11; Conservative 4; Mismatches 7; Indels 4; Gaps 1;

OY 8 YHNVEMQWTTCKPETTNCVPOERE 33  
 Db 11 HHLOV-----TSCPDATTCTDQDK 32

RESULT 9  
 E86464  
 hypothetical protein F12G12.12 [imported] - *Arabidopsis thaliana*  
 C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 C/Accession: E86464  
 R/Theologian, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: E86464  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-463 <STO>  
 A/Cross-references: GB:AE005172; NID:g10086471; PIDN:AA612531.1; GSPDB:GN00141  
 C/Genetics: 1  
 A/Map position: 1

Query Match 28.6%; Score 54; DB 2; Length 463;  
 Best Local Similarity 58.3%; Pred. No. 12;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 16 WTTCKPETTNC 27  
 Db 392 WVKCERPDPTNC 403

RESULT 10  
 A47142  
 cysstatin D precursor - human  
 C/Species: *Homo sapiens* (man)  
 C/Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 16-Jul-1999  
 C/Accession: A47142; S18212  
 R/Freije, J.P.; Balbin, M.; Abrahamson, M.; Velasco, G.; Dalboge, H.; Grubb, A.; Lopez-C J.; Biol. Chem. 268, 15737-15744, 1993  
 A/Title: Human cysstatin D. cDNA cloning, characterization of the *Escherichia coli* expres  
 A/Reference number: A47142; MUID:93340179; PMID:8340398  
 A/Accession: A47142  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-142 <PRE>

A/Cross-references: GB:X70377; NID:g398710; PIDN:CAA49838.1; PID:g398711  
 A/Note: single residue difference between this report and S18218 was investigated and shk  
 R/Freije, J.P.; Abrahamson, M.; Olafsson, I.; Velasco, G.; Grubb, A.; Lopez-Otin, C.  
 J. Biol. Chem. 266, 20538-20543, 1991  
 A/Title: Structure and expression of the gene encoding cysstatin D, a novel human cysteine  
 A/Reference number: S18212; MUID:92041895; PMID:1939105  
 A/Accession: S18212  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-45 'C', 47-142 <PRE>  
 A/Cross-references: EMBL:X59964; NID:g30263; PIDN:CAA42590.1; PID:g30264  
 A/Genetics: 3  
 A/Map position: 20p11.21-20p11.21  
 A/Introns: 77/3; 115/3  
 C/Superfamily: cysstatin; cysstatin homology  
 C/Keywords: cysteine proteinase inhibitor; extracellular protein; saliva  
 F/1-20/Domain: signal sequence #status predicted <SIG>  
 F/30-142/Domain: cysstatin homology <CYS>

Query Match 28.0%; Score 53; DB 2; Length 142;  
 Best Local Similarity 37.9%; Pred. No. 5;  
 Matches 11; Conservative 4; Mismatches 12; Indels 2; Gaps 1;

OY 1 QVTDHLEHNVEMQWTTCKP--PETTNC 27  
 Db 77 QVGVGVVYFNVKFGRTCTKSGPNLNC 105

RESULT 11  
 JH0563  
 metabotropic glutamate receptor 4 precursor - rat  
 C/Species: *Rattus norvegicus* (Norway rat)  
 C/Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 17-Nov-2000  
 C/Accession: JH0563; I58149  
 R/Tanabe, Y.; Mausu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.  
 Neuron 8, 169-179, 1992  
 A/Title: A family of metabotropic glutamate receptors.  
 A/Reference number: JH0561; MUID:92110002; PMID:1309649  
 A/Accession: JH0563  
 A/Molecule type: mRNA  
 A/Residues: 1-912 <TRAN>  
 A/Experimental source: Brain  
 R/O'Hara, P.J.; Sheppard, P.O.; Thøgersen, H.; Venezia, D.; Haldeman, B.A.; McGrane, V.; Neuron 11, 41-52, 1993  
 A/Title: The ligand-binding domain in metabotropic glutamate receptors is related to bact  
 A/Reference number: I58149; MUID:9332699; PMID:8338667  
 A/Accession: I58149  
 A/Status: preliminary; translated from GB/EMBL/DDAJ  
 A/Molecule type: mRNA  
 A/Residues: 1-123 'R', 125-912 <RES>  
 A/Cross-references: GB:M90518; NID:g205400; PIDN:AA93190.1; PID:g205401  
 C/Comment: This protein is coupled to a G protein and evokes a variety of functions by me  
 A/Genetics: 4  
 A/Map position: 20p11.21-20p11.21  
 C/Superfamily: metabotropic glutamate receptor 4  
 C/Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phospho  
 F/1-32/Domain: signal sequence #status predicted <SIG>  
 F/33-912/Domain: metabotropic glutamate receptor 4 #status predicted <MET>  
 F/625-645/Domain: transmembrane #status predicted <TRI>  
 F/657-675/Domain: transmembrane #status predicted <TRI>  
 F/700-720/Domain: transmembrane #status predicted <TRI>  
 F/751-772/Domain: transmembrane #status predicted <TRI>  
 F/786-807/Domain: transmembrane #status predicted <TRI>  
 F/822-847/Domain: transmembrane #status predicted <TRI>  
 F/98-301,454,484,569/Binding site: carbohydrate (Aen) (covalent) #status predicted  
 F/621,669,695,859,870/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 28.0%; Score 53; DB 2; Length 912;  
 Best Local Similarity 39.0%; Pred. No. 34;  
 Matches 16; Conservative 3; Mismatches 6; Indels 16; Gaps 4;

QY 3 TDHLEHYLWVE-MOW-----TTCQKPEITNCVQERE 33  
DB 496 TDHLE-HLRTERMQWPGSGQLPRISCSLP---CQGERKK 530

RESULT 12  
T29165  
hypothetical protein C01B7.6 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C/Accession: T29165

R/Moesener, J.; Bradshaw, H.

submitted to the EMBL Data Library, March 1996

A/Description: The sequence of C. elegans cosmid C01B7.

A/Reference number: 220581

A/Accession: T29165

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-3766 <MOB>

A/Cross-references: EMBL:U53147; PIDN:AAA96117.1; GSPDB:GN00023; CBSP:C01B7.6

A/Experimental source: strain Bristol N2; clone C01B7

C/Genetics:

A/Map position: 5

A/Introns: 63/1; 137/3; 282/2; 430/3; 479/1; 1743/3; 2424/3; 2606/3; 2926/3; 2985/3; 321

Query Match 28.0%; Score 53; DB 2; Length 3766;

Best Local Similarity 36.4%; Pred. No. 1.5e+02;

Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 5 HLEHYLWVEWQWTTCKPEITN 26

DB 2415 YLRFHQLNHNWQSCRPSCN 2436

RESULT 13

A36163

cystatin C precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 16-Jul-1999

C/Accession: A36163

R/Solem, M.; Rawson, C.; Lindburg, K.; Barnes, D.

Biochem. Biophys. Res. Commun. 172, 945-951, 1990

A/Title: Transforming growth factor beta regulates cystatin C in serum-free mouse embryo

A/Reference number: A36163; MUID:91054522; PMID:2241963

A/Accession: A36163

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-140 <SOL>

A/Cross-references: EMBL:M59470; NID:G192911; PIDN:AAA63298.1; PID:G192912

C/Superfamily: cystatin; cystatin homology

F/29-140/Domain: cystatin homology <CYS>

F/93-103,117-137/Disulfide bonds: #status predicted

Query Match 27.5%; Score 52; DB 2; Length 140;

Best Local Similarity 41.4%; Pred. No. 6.8;

Matches 12; Conservative 5; Mismatches 10; Indels 2; Gaps 1;

QY 1 QVTDHLEHYLWVEWQWTTCKPEITN 27

DB 75 QLVAGVYFPDVENGRITTKSQTNLTDC 103

RESULT 14

B96614

hypothetical protein T15M6.10 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C/Accession: B96614

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: B96614

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-371 <STO>

A/Cross-references: GB:AE005173; NID:G11067306; PIDN:AA28833.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

A/Suprafamily: Arabidopsis thaliana hypothetical protein F24M12.210

Query Match 27.2%; Score 51.5; DB 2; Length 371;

Best Local Similarity 29.7%; Pred. No. 22;

Matches 11; Conservative 3; Mismatches 10; Indels 13; Gaps 1;

QY 1 QVTDHLEHYLWVEWQWTTCKPEITN 24

DB 177 KYTDHTRFHEKPELMDSGRSLNGLWYATVNSPHT 213

RESULT 15

UDBO

cystatin - bovine

N/Alternate names: thiol proteinase inhibitor

C/Species: Bos primigenius taurus (cattle)

C/Date: 28-Feb-1986 #sequence\_revision 28-Feb-1986 #text\_change 06-Dec-1996

C/Accession: A01271

R/Hirado, M.; Tsunagaawa, S.; Sakiyama, F.; Nishibe, M.; Fujii, S.

FEBS Lett. 186, 41-45, 1985

A/Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase inh

A/Reference number: A01271; MUID:85231205; PMID:3891407

A/Accession: A01271

A/Molecule type: protein

A/Residues: 1-112 <HR>

C/Superfamily: cystatin; cystatin homology

C/Keywords: colostrum; cysteine proteinase inhibitor

F/2-112/Domain: cystatin homology <CYS>

F/46-52/Region: inhibitory #status predicted

F/66-76,90-110/Disulfide bonds: #status predicted

Query Match 26.5%; Score 50; DB 1; Length 112;

Best Local Similarity 43.5%; Pred. No. 10;

Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 QVTDHLEHYLWVEWQWTTCKPE 23

DB 48 QVSGMNYFLDVELGRITTKSQ 70

Search completed: March 18, 2004, 14:23:17  
Job time: 7.35119 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 3.92857 seconds  
(without alignments)  
437.389 Million cell updates/sec

Title: US-09-941-314-12  
Perfect score: 189  
Sequence: 1 QVTDHLEHNLNEMQWTCCKPRTNCVPOERE 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	137	CS11_HUMAN	Q9H112 homo sapien
2	104	55.0	139	CS11_MOUSE	Q94269 mus musculu
3	95	50.3	142	CST8_RAT	O88969 rattus norv
4	89	47.1	142	CST8_MOUSE	P22766 mus musculu
5	70	37.0	142	CST8_HUMAN	O60676 homo sapien
6	65	34.4	127	CYT8_RAT	P14841 rattus norv
7	63	33.4	165	CSTL_HUMAN	Q9H114 homo sapien
8	57	30.2	146	CYT8_SAISC	O39093 hamstri sci
9	56	29.6	140	CYT8_MOUSE	P14460 mus musculu
10	55	29.1	4540	DYHC_PARTE	Q27171 paramecium
11	54.5	28.8	111	CYT8_BIRAR	P08935 bitis ariet
12	54	28.6	146	CYT8_HUMAN	P01034 homo sapien
13	54	28.6	146	CYT8_MACMU	O19092 macaca mula
14	53	28.0	142	CYT8_HUMAN	P83323 homo sapien
15	53	28.0	912	MGR4_RAT	P14323 rattus norv
16	52	27.5	148	CYT8_RABIT	O97862 cyttocolagus
17	50	26.5	148	CYT8_BOVIN	P01033 bos taurus
18	50	26.5	632	Y252_DROME	P84066 homo sapien
19	50	26.5	641	Y252_DROME	Q94545 drosophila
20	50	26.5	716	GFAL1_YEAST	P14742 saccharomyc
21	50	26.5	801	PRD4_HUMAN	O94058 homo sapien
22	49	25.9	107	TIM4_BOVIN	O97563 bos taurus
23	49	25.9	145	CYT8_HUMAN	O76096 homo sapien
24	49	25.9	170	TIM4_RABIT	O97591 cyttocolagus
25	49	25.9	224	TIM4_HUMAN	O99727 homo sapien
26	49	25.9	328	SGS3_DROER	P13730 drosophila
27	49	25.9	875	NPP3_HUMAN	O14638 h ecomocula
28	49	25.9	1276	GIL1_ENTRI	P32022 entamoeba h
29	48.5	25.7	144	CYT8_MOUSE	O89058 mus musculu
30	48	25.4	455	PEL6_ARATH	O64510 arabidopsis
31	48	25.4	912	MGR4_HUMAN	O14833 homo sapien
32	47.5	25.1	214	PUR3_YEAST	P04161 saccharomyc
33	47	24.9	456	PUR3_BUCAL	P57351 buchiera ap

34	47	24.9	541	1	ACEA_YARLI	P41555 yarrowia li
35	46.5	24.6	239	1	PDJ1_ANASP	O82080 anabena sp
36	46.5	24.6	712	1	GFAL1_CANAL	P53704 candida alb
37	46.5	24.6	768	1	ENV_SIVAL	P27757 simian immu
38	46	24.3	106	1	ILB1_CAREL	O09626 caenorhabdi
39	46	24.3	367	1	HIS6_SYNY3	P73807 synechocyst
40	46	24.3	511	1	RNE_FORPU	P51211 porphyra pu
41	46	24.3	548	1	CUI2_SCHPO	O59755 schistosach
42	46	24.3	676	1	SYM_ECO57	O8x767 escherichia
43	46	24.3	676	1	SYM_ECOL6	O8fx8 escherichia
44	46	24.3	676	1	SYM_ECOL1	P00959 escherichia
45	46	24.3	759	1	PEPX_STRGN	Q93m42 streptococc

## ALIGNMENTS

RESULT 1  
ID CS11\_HUMAN STANDARD; PRT; 137 AA.  
AC Q9H112; Q9H113; 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin II precursor.  
GN CstII OR CST8L.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhand P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Frazer A.A., French L., Garner P.,  
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehaeslath M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurtry A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
RA Rice C.M., Rose M.T., Scott C.B., Sehra H.K., Showkeen R., Sims S.,  
RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulten J.B.,  
RA Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q9H112-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q9H112-2; Sequence=VSP\_001260;  
CC Note=No experimental confirmation available;  
CC -1- SIMILARITY: Belongs to the cystatin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AL096677; CAC13170.1; -  
 DR EMBL; AL096677; CAC17423.1; -  
 DR HSSP; P01038; 1A90.  
 DR Genew; HGNC:15959; CST11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; Cy; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
 DR Thiol protease inhibitor; Signal; Alternative splicing.  
 DR SIGNAL 1 25  
 FT CHAIN 26 137  
 FT SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 93 101 BY SIMILARITY.  
 FT DISULFID 114 134 BY SIMILARITY.  
 FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARSPLIC 76 110 Missing (in isoform 2).  
 FT /FTID=VSP 001260.  
 SQ SEQUENCE 137 AA; 16375 MM; C5856C39A585C3B CRC64;

Query Match 100.0%; Score 189; DB 1; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-19;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33  
 Db 75 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 107

RESULT 2  
 CS11\_MOUSE  
 ID CS11\_MOUSE STANDARD; PRT; 139 AA.  
 AC Q9D269;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin 11 precursor.  
 GN CST11.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada S.,  
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Rietshamn W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Brownstein S., Hill D., Hofmann M., Hume D.A., Kamitani M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald C., Seya T., Shibata Y., Storch K.-F.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AK020300; BAB32061.1; -  
 DR HSSP; P01034; 1G96.  
 DR MGD; MGI:1925490; Cst11.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; Cy; 1.  
 DR PROSITE; PS00287; CYSTATIN; FALSE\_NEG.  
 DR Thiol protease inhibitor; Signal.  
 DR SIGNAL 1 28  
 FT CHAIN 29 139  
 FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 94 102 BY SIMILARITY.  
 FT DISULFID 115 135 BY SIMILARITY.  
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT SEQUENCE 139 AA; 16217 MM; F228D9815FA32640 CRC64;

Query Match 55.0%; Score 104; DB 1; Length 139;  
 Best Local Similarity 63.6%; Pred. No. 1.3e-07;  
 Matches 21; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33  
 Db 76 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 108

RESULT 3  
 CST8\_RAT  
 ID CST8\_RAT STANDARD; PRT; 142 AA.  
 AC O88969;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin  
 DE 8).  
 GN CST8 OR CRES.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Epididymis;  
 RX MEDLINE=99247899; PubMed=10229662;  
 RA Cornwall G.A., Heia N., Sutton H.G.;  
 RA "Structure, alternative splicing and chromosomal localization of the  
 RT cystatin-related epididymal spermatogenic gene";  
 RL Biochem. J. 340:85-93(1999).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF090692; AAC36317.1; -  
 DR HSSP; P01034; 1G96.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; Cy; 1.

KW Thiol protease inhibitor; Signal.  
 PT SIGNAL 1 19 POTENTIAL.  
 PT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 PT SITE 77 81 PROTEIN.  
 FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 119 139 BY SIMILARITY.  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 142 AA; 16246 MW; FB873FAA6BCAB34 CRC64;  
 Query Match 50.3%; Score 95; DB 1; Length 142; -  
 Best Local Similarity 45.5%; Pred. No. 2,4e-06;  
 Matches 15; Conservative 10; Mismatches 6; Indels 2; Gaps 1;  
 QY 1 QVTDHLEHYLVNEMQWTTCKP--ETTNCVPOE 31  
 77 QITDRMEHYDVOISRSNCKPLNTENCIPQK 109  
 Db  
 RESULT 4  
 CST8\_MOUSE STANDARD; PRT; 142 AA.  
 ID CST8\_MOUSE 089102;  
 AC P32766; 089102;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-  
 DE related epididymal specific protein) (Cystatin 8).  
 GN CST8 OR CRIS.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H, and CD-1;  
 RX MEDLINE=99247899; PubMed=10229662;  
 RA Cornwall G.A., Hala N., Sutton H.G.;  
 RT "Structure, alternative splicing and chromosomal localization of the  
 RT cystatin-related epididymal spermatogenic gene.";  
 RL Biochem. J. 340:85-93(1999).  
 RN [2]  
 RP SEQUENCE OF 4-142 FROM N.A.  
 RC TISSUE=Epididymis;  
 RX MEDLINE=93078799; PubMed=1280328;  
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;  
 RT "The CRIS gene: a unique testis-regulated gene related to the cystatin  
 RT family is highly restricted in its expression to the proximal region  
 RT of the mouse epididymis.";  
 RL Mol. Endocrinol. 6:1653-1664(1992).  
 RN [1]  
 RP FUNCTION: Performs a specialized role during sperm development and  
 RP maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epithelium.  
 CC -1- INDUCTION: Testicular factors or hormones other than androgens  
 CC present in the testicular fluid may be involved in the regulation  
 CC of CRIS gene expression.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF091503; AAC61754.1; -  
 CC EMBL; AF090691; AAC6316.1; -  
 CC EMBL; S49926; AAC35390.1; -

DR PIR; A45361; A45361.  
 DR HSSP; P01034; 1G96.  
 DR MGD; MGI:107161; Cst8.  
 DR Interpro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 KW Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 FT SITE 77 81 PROTEIN.  
 FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 119 139 BY SIMILARITY.  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 4 15 PLMSLFLFTIP --GTRDEQVGSOK (IN REF. 2).  
 SQ SEQUENCE 142 AA; 16288 MW; 50B446B98F6673E CRC64;  
 Query Match 47.1%; Score 89; DB 1; Length 142;  
 Best Local Similarity 41.2%; Pred. No. 1,7e-05;  
 Matches 14; Conservative 11; Mismatches 7; Indels 2; Gaps 1;  
 QY 1 QVTDHLEHYLVNEMQWTTCKP--ETTNCVPOE 32  
 77 QITDRMEHYDVOISRSNCKPLNTENCIPQK 110  
 Db  
 RESULT 5  
 CST8\_HUMAN STANDARD; PRT; 142 AA.  
 ID CST8\_HUMAN 060676;  
 AC 060676;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin  
 DE 8).  
 GN CST8 OR CRIS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=95344753; PubMed=7619504;  
 RA Cornwall G.A., Hann S.R.;  
 RT "Transient appearance of CRIS protein during spermatogenesis and  
 RT caput epididymal sperm maturation.";  
 RL Mol. Reprod. Dev. 41:37-46(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.B., Bridgman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dham P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grainger D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lelevasilho M.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McCormack L.J., McHay K., McIlrath T.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Pratchallam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
 RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.B.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.D., Whitaker P., Willey D., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.,  
 RL "The DNA sequence and comparative analysis of human chromosome 20.",  
 CC Nature 414:865-871(2001).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epithelium.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL, AF059244; AAC14707.1; -  
 CC EMBL, AL109954; CAB64234.1; -  
 CC HSSP, P01034; 1G96.  
 CC Genew, HGNC:2480; CST8.  
 CC GO, GO:0004669; F:cysteine protease inhibitor activity; TAS.  
 CC InterPro: IPR000010; Cystatin.  
 CC Pfam, PF00031; cystatin; 1.  
 CC SMART, SM00043; CY; 1.  
 CC K101 protease inhibitor; Signal; Polymorphism.  
 CC FT SIGNAL 1 21 POTENTIAL.  
 CC FT CHAIN 22 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC  
 CC PROTEIN.  
 CC FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).  
 CC FT DISULFID 95 105 BY SIMILARITY.  
 CC FT DISULFID 119 139 BY SIMILARITY.  
 CC FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT VARIANT 142 142 A -> P (in dbSNP:1054633).  
 CC /FTID=VAR\_014527.  
 CC FT SEQUENCE 142 AA; 16275 MW; 9A3512757E0F4ECD CRC64;  
 CC SQ  
 CC Query Match 37.0%; Score 70; DB 1; Length 142;  
 CC Best Local Similarity 48.5%; Pred. No. 0.0076;  
 CC Matches 16; Conservative 7; Mismatches 8; Indels 2; Gaps 1;  
 CC  
 CC QY 1 QVTDHLEVHNLVEMQWTCCKPFTN--CVPOE 31  
 CC Db 77 QVTDHLEVHNLVEMQWTCCKPFTN--CVPOE 109  
 CC  
 CC RESULT 6  
 CC CYTC\_RAT STANDARD; PRT; 127 AA.  
 CC ID CYTC\_RAT STANDARD; PRT; 127 AA.  
 CC AC P14841;  
 CC DT 01-APR-1990 (Rel. 14, Created)  
 CC DT 01-APR-1990 (Rel. 14, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Cystatin C precursor (Fragment).  
 CC GN CST3.  
 CC OS Rattus norvegicus (Rat).  
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CC OX NCBI\_TaxID=10116;  
 CC RN  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=Buffalo;  
 CC RX MEDLINE=90092122; PubMed=2689174;  
 CC RA Cole T., Dickson P.W., Esmard F., Averill F., Risbridger G.,  
 RA Gauthier F., Schreiber G.,  
 RT "The CDNA structure and expression analysis of the genes for the

RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin  
 RT in rat brain.";  
 RL Eur. J. Biochem. 186:35-42(1989).  
 RN  
 RN [2]  
 RN SEQUENCE OF 8-127.  
 RX MEDLINE=90380276; PubMed=2400577;  
 RA Bernard F., Bernard A., Faucher D., Capony J.-P., Derancourt J.,  
 RA Brillard M., Gauthier F.;  
 RT "Rat cystatin C: the complete amino acid sequence reveals a site for  
 RT N-glycosylation.";  
 RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).  
 RN  
 RN [3]  
 RN SEQUENCE OF 8-49.  
 RX MEDLINE=88313020; PubMed=3044831;  
 RA Bernard A., Bernard F., Faucher D., Gauthier F.;  
 RT "Two rat homologues of human cystatin C.";  
 RL FEBS Lett. 236:475-478(1988).  
 RN  
 RN [4]  
 RN SEQUENCE OF 8-20.  
 RC TISSUE=Semtoli cells;  
 RX MEDLINE=92225121; PubMed=1563513;  
 RA Bernard A., Bernard F., Guillon F., Gauthier F.;  
 RT "Production of the cysteine proteinase inhibitor cystatin C by rat  
 RT Semtoli cells.";  
 RL FEBS Lett. 300:131-135(1992).  
 CC  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 CC thought to serve an important physiological role as a local  
 CC regulator of this enzyme activity. Known to inhibit cathepsin B,  
 CC H. and L.  
 CC  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL, X16957; CAA34831.1; -  
 CC PIR, S07085; S07085.  
 CC PIR, S10587; S10587.  
 CC HSSP, P01034; 1G96.  
 CC InterPro: IPR000010; Cystatin.  
 CC Pfam, PF00031; cystatin; 1.  
 CC SMART, SM00043; CY; 1.  
 CC DR PROSITE, PS00287; CYSTATIN; 1.  
 CC K101 protease inhibitor; Signal.  
 CC FT NON\_TER 1 1  
 CC FT SIGNAL 1 7  
 CC FT CHAIN 8 127 CYSTATIN C.  
 CC FT ACT SITE 18 18 REACTIVE SITE.  
 CC FT SITE 62 62 SECONDARY AREA OF CONTACT.  
 CC FT DISULFID 80 90 BY SIMILARITY.  
 CC FT DISULFID 104 124 BY SIMILARITY.  
 CC FT CONFLICT 25 25 A -> E (IN REF. 2).  
 CC FT SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;  
 CC SQ  
 CC Query Match 34.4%; Score 65; DB 1; Length 127;  
 CC Best Local Similarity 48.3%; Pred. No. 0.034;  
 CC Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;  
 CC  
 CC QY 1 QVTDHLEVHNLVEMQWTCCKPFTN--TNC 27  
 CC Db 62 QVTDHLEVHNLVEMQWTCCKPFTN--TNC 90  
 CC  
 CC RESULT 7  
 CC CSTL\_HUMAN STANDARD; PRT; 165 AA.  
 CC ID CSTL\_HUMAN STANDARD; PRT; 165 AA.  
 CC AC Q9H114;  
 CC DT 28-FEB-2003 (Rel. 41, Created)  
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)



OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=91054522; PubMed=2241983;  
 RA Solem M., Rawson C., Lindburg K., Barnes D.;  
 RT "Transforming growth factor beta regulates cystatin C in serum-free  
 mouse embryo (SEME) cells."  
 RL Biochem. Biophys. Res. Commun. 172:945-951(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv; TISSUE=liver;  
 RX MEDLINE=95137392; PubMed=7835704;  
 RA Huh C., Nagle J.W., Kozak C.A., Abrahamson M., Karlsson S.;  
 RT "Structural organization, expression and chromosomal mapping of the  
 mouse cystatin-C-coding gene (Cst3).";  
 RL Gene 152:121-126(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ILS, and ISS;  
 RX MEDLINE=21363810; PubMed=11471062;  
 RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Camille J.,  
 RA Beeson M., Gordon L., Bennett B., Johnson T.E., Skeja J.M.;  
 RT "High-throughput sequence identification of gene coding variants  
 within alcohol-related QTLs";  
 RL Mamm. Genome 12:657-663(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Butler K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueidi T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,  
 RA Vallation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Foley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,  
 RA Blakeley R.C., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield A.S., Krzywinski M.I., Skalska J., Smalhe D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 CC thought to serve an important physiological role as a local  
 CC regulator of this enzyme activity.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M59470; AAA63298.1; -;  
 CC EMBL; U10098; AAB41056.1; -;  
 CC EMBL; AF483486; AAL90760.1; -;  
 CC EMBL; AF483487; AAL90761.1; -;  
 CC EMBL; BC002072; AA02072.1; -;  
 CC PIR; A36163; A36163.  
 CC HSSP; P01034; 1G96.  
 CC MGD; MGI:102519; Cst3.  
 CC Interpro: IPR000010; Cystatin.  
 CC Pfam; PF00031; Cystatin; 1.  
 CC SMART; SM00043; Cy; 1.

DR PROSITE; PS00287; CYSTATIN; 1.  
 KW Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 140 CYSTATIN C.  
 FT ACT SITE 31 31 REACTIVE SITE.  
 FT SITE 75 79 SECONDARY AREA OF CONTACT.  
 FT DISULFID 93 103 BY SIMILARITY.  
 FT DISULFID 117 137 BY SIMILARITY.  
 FT CONFLICT 16 16 A -> G (IN REF. 1).  
 FT CONFLICT 84 84 L -> F (IN REF. 1).  
 SQ SEQUENCE 140 AA; 15531 MW; 3A563406D58D0F5 CRC64;  
 QY Query Match 29.6%; Score 56; DB 1; Length 140;  
 DB Best Local Similarity 44.8%; Pred. No. 0.68;  
 DB Matches 13; Conservative 5; Mismatches 9; Indels 2; Gaps 1;  
 QY 1 QVTDHLEHYHNVEMQWTCCKPPT--TNC 27  
 DB 75 QLVAGVNYFLDVEWGRITCKSQNTLTDG 103  
 RESULT 10  
 DYHC PARTE STANDARD; PRT; 4540 AA.  
 ID Q27171;  
 DC 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Dynein heavy chain, cytosolic (DYHC).  
 GN DHC-8.  
 OS Paramaecium tetraurelia.  
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;  
 OC Paramacium.  
 OX NCBI\_TaxID=5888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Stock 51;  
 RX MEDLINE=96157890; PubMed=8589455;  
 RA Kandi K.A., Forney J.D., Asai D.J.;  
 RT "The dynein genes of Paramaecium tetraurelia: the structure and  
 RT expression of the ciliary beta and cytoplasmic heavy chains";  
 RL Mol. Biol. Cell 6:1549-1562(1995).  
 CC -1- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular  
 CC retrograde motility of vesicles and organelles along microtubules.  
 CC Dynein has ATPase activity; the force-producing power stroke is  
 CC thought to occur on release of ADP.  
 CC -1- SUBUNIT: Consists of at least two heavy chains and a number of  
 CC intermediate and light chains.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem  
 CC (which binds cargo and interacts with other dynein components),  
 CC and the head or motor domain. The motor contains six tandemly-  
 CC linked AAA domains in the head, which form a ring. A stalk-like  
 CC structure (formed by two of the coiled coil domains) protrudes  
 CC between AAA 4 and AAA 5 and terminates in a microtubule-binding  
 CC site. A seventh domain may also contribute to this ring; it is not  
 CC clear whether the N-terminus or the C-terminus forms this extra  
 CC domain. There are four well-conserved and two non-conserved ATPase  
 CC sites, one per AAA domain. Probably only one of these (within AAA  
 CC 1) actually hydrolyzes ATP, the others may serve a regulatory  
 CC function.  
 CC -1- SIMILARITY: Belongs to the dynein heavy chain family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U20449; AAA75445.1; -;  
 CC PIR; T30838; T30838.

```

DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR004273; Dyncin_heavy.
DR Pfam: PF03028; Dyncin_heavy; 1.
DR SMART: SM00382; AAA; 3.
KW Motor protein; Dyncin; Microtubule; ATP-binding; Repeat; Coiled coil.
FT DOMAIN 1 1796 STEM (BY SIMILARITY).
FT DOMAIN 1 1797 2018 AAA 1 (BY SIMILARITY).
FT DOMAIN 2091 2348 AAA 2 (BY SIMILARITY).
FT DOMAIN 2457 2705 AAA 3 (BY SIMILARITY).
FT DOMAIN 2796 3056 AAA 4 (BY SIMILARITY).
FT DOMAIN 3076 3367 STALK (BY SIMILARITY).
FT DOMAIN 3444 3673 AAA 5 (BY SIMILARITY).
FT DOMAIN 3908 4123 AAA 6 (BY SIMILARITY).
FT DOMAIN 440 482 COILED COIL (POTENTIAL).
FT DOMAIN 698 722 COILED COIL (POTENTIAL).
FT DOMAIN 794 827 COILED COIL (POTENTIAL).
FT DOMAIN 975 995 COILED COIL (POTENTIAL).
FT DOMAIN 1169 1251 COILED COIL (POTENTIAL).
FT DOMAIN 1295 1311 COILED COIL (POTENTIAL).
FT DOMAIN 3076 3182 COILED COIL (POTENTIAL).
FT DOMAIN 3289 3367 COILED COIL (POTENTIAL).
FT DOMAIN 3653 3688 COILED COIL (POTENTIAL).
FT DOMAIN 3820 3851 COILED COIL (POTENTIAL).
FT DOMAIN 4238 4259 COILED COIL (POTENTIAL).
FT DOMAIN 4313 4342 COILED COIL (POTENTIAL).
FT NP_BIND 1835 1842 ATP (POTENTIAL).
FT NP_BIND 2129 2136 ATP (POTENTIAL).
FT NP_BIND 2496 2503 ATP (POTENTIAL).
FT NP_BIND 2834 2841 ATP (POTENTIAL).
SQ SEQUENCE 4540 AA; 528626 MW; 0C6103148BF895F9 CRC64;

Query Match 29.1%; Score 55; DB 1; Length 4540;
Best Local Similarity 40.6%; Pred. No. 34;
Matches 13; Conservative 5; Mismatches 8; Indels 6; Gaps 2;

7 EYHNL-----VEMQWTTCKPFTTCVPORE 33
Db 1191 ERYLNOQIOEIEFQWMT-SKPDSDCSPNFAE 1221

RESULT 11
CYT BITAR STANDARD; PRT; 111 AA.
AC P08935;
AT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DB Cystatin.
OS Bitis arietans (African puff adder).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Viperinae; Bitis.
ON NCB1_Taxid=8692;
OX
RN
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=88076861; PubMed=3500714;
RA Ritonja A., Evans H.J., Machleidt W., Barrett A.J.;
RT "Amino acid sequence of a cystatin from venom of the African puff
addor (Bitis arietans).";
RL Biochem. J. 246:799-802(1987).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the cystatin family.
DR HSSP; P01038; 1CEW.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART; SM00043; CY 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor.
FT ACT SITE 3 REACTIVE SITE.
FT SITE 47 51 SECONDARY AREA OF CONTACT.
FT DISUFID 65 81 PROBABLE.

```

```

FT VARIANT 57 57 M -> T (IN EQUAL AMOUNT).
SQ SEQUENCE 111 AA; 12678 MW; 1A31B6B246ACAL0C CRC64;

Query Match 28.8%; Score 54.5; DB 1; Length 111;
Best Local Similarity 33.3%; Pred. No. 0.87;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

1 QVTDHLEHNLVEMQWTTCKPFTTCVPORE 33
Db 47 QVVSQVRYLMEHLKTKTKVGRPKGVQIQCNIPLPBNQ 88

RESULT 12
CYTC_HUMAN STANDARD; PRT; 146 AA.
AC P01034;
AT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)
DE (Post-gamma-globulin).
GN CST3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_Taxid=9606;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=87219149; PubMed=3495457;
RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;
RT "Molecular cloning and sequence analysis of cDNA coding for the
precursor of the human cysteine proteinase inhibitor cystatin C.";
RL FEBS Lett. 216:229-233(1987).
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RX MEDLINE=90303202; PubMed=2363674;
RA Abrahamson M., Olafsson I., Palsdottir A., Ulvsbaeck M., Lundwall A.,
RA Jensen O., Grubb A.;
RT "Structure and expression of the human cystatin C gene.";
RL Biochem. J. 268:287-294(1990).
RN
RP SEQUENCE FROM N.A. (HCHWA VARIANT).
RC TISSUE=Brain;
RX MEDLINE=89235594; PubMed=3541223;
RA Levy E., Lopez-Otin C., Ghiso J., Gellner D., Frangione B.;
RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is
related to a mutation in the cystatin C gene, an inhibitor of
cysteine proteases.";
RL J. Exp. Med. 169:1771-1778(1989).
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=89350949; PubMed=2764935;
RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.;
RA Iemura S., Sanada K.;
RT "The human cystatin C gene (CST3) is a member of the cystatin gene
family which is localized on chromosome 20.";
RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).
RN
RP SEQUENCE FROM N.A.
RA Dickinson D.P., Hewett-Emmett D., Thiasse M.;
RT "Acquisition of complex patterns of differential expression in
epithelial cell populations during the evolution of type 2 cystatin
genes.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Bease D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,

```

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clapp M., Clark R.E., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier G.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Gratlam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jerosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Levesajaho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurtry A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Rameay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
 RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W., Thorpe A.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmberg L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:665-871(2001).  
 [17]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richerds S., Wolter K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Folley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [18]  
 RN SEQUENCE OF 27-146.  
 RX MEDLINE=8222268; PubMed=6283552;  
 RA Grubb A., Loeffberg H.;  
 RT "Human gamma-trace, a basic microprotein: amino acid sequence and  
 presence in the adenylophosphates";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027(1982).  
 [19]  
 RN SEQUENCE OF 27-73.  
 RX MEDLINE=84110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longner M., Riconja A., Eropkin M., Borchart U.,  
 RA Mechleide W.;  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 of cystatin from chicken egg white";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).  
 [10]  
 RN SEQUENCE OF 27-76.  
 RX MEDLINE=84128015; PubMed=6365094;  
 RA Brzin J., Popovic T., Turk V.;  
 RT "Human cystatin, a new protein inhibitor of cysteine proteinases";  
 RL Biochem. Biophys. Res. Commun. 118:103-109(1984).  
 [11]  
 RN DISULFIDE BONDS.  
 RA Grubb A., Loeffberg H., Barrett A.J.;  
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
 cystatin.";

RL FEBS Lett. 170:370-374(1984).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.  
 RX MEDLINE=21173909; PubMed=11276250;  
 RA Janowski R., Kozak M., Jankowska E., Grzonka Z., Grubb A.,  
 RA Abramson M., Jaskolski M.;  
 RT "Human cystatin C, an amyloidogenic protein, dimerizes through  
 three-dimensional domain swapping";  
 RL Nat. Struct. Biol. 8:316-320(2001).  
 RN [13]  
 RP VARIANT GLN-94.  
 RX MEDLINE=92316504; PubMed=1352269;  
 RA Abramson M., Jankowska E., Olafsson I., Jensen O., Grubb A.;  
 RT "Hereditary cystatin C amyloid angiopathy: identification of the  
 disease-causing mutation and specific diagnosis by polymerase chain  
 reaction based analysis";  
 RL Hum. Genet. 89:377-380(1992).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 thought to serve an important physiological role as a local  
 regulator of this enzyme activity.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,  
 vas deferens, brain, thymus, and ovary and the lowest in the  
 submandibular gland.  
 CC -1- DISEASE: Defects in CST3 are a cause of hereditary cerebral  
 hemorrhage with amyloidosis (HCHWA) (MIM:105150); also known as  
 cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis  
 Icelandic type. HCHWA is characterized by a thickening of the  
 cerebral arteries walls with deposition of material with the  
 characteristics of amyloid.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X05607; CAA29096.1; -;  
 DR EMBL: X52255; CAA36497.1; -;  
 DR EMBL: M27891; AAA52164.1; -;  
 DR EMBL: M27889; AAA52164.1; JOINED.  
 DR EMBL: M27890; AAA52164.1; JOINED.  
 DR EMBL: X61681; CAA43856.2; -;  
 DR EMBL: X61682; CAA43856.2; JOINED.  
 DR EMBL: X61683; CAA43856.2; JOINED.  
 DR EMBL: AF319564; AAK11570.1; -;  
 DR EMBL: AL121894; CAC05424.1; -;  
 DR EMBL: BC013083; AAI13083.1; -;  
 DR PIR: S10216; UDHU.  
 DR PDB: 1G96; 06-APR-01.  
 DR Genew: HGNC:2475; CST3.  
 DR MIM: 604312; -;  
 DR MIM: 105150; -;  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin, 1.  
 DR SMART: SM00043; CY. 1.  
 DR PROSITE: PS00287; CYSTATIN, 1.  
 DR Thiol protease inhibitor; Amyloid; signal; Disease mutation;  
 KW Polymorphism; 3D-structure.  
 FT SIGNAL 1 26  
 FT CHAIN 27 146 CYSTATIN C.  
 FT ACT\_SITE 37 37 REACTIVE SITE.  
 FT SITE 81 85 SECONDARY AREA OF CONTACT.  
 FT DISULFID 99 109  
 FT DISULFID 123 143  
 Query Match 28.6%; Score 54; DB 1; Length 146;  
 Best Local Similarity 41.4%; Pred. No. 1.3; Indels 2; Gaps 1;  
 Matches 12; Conservative 4; Mismatches 11;



```

DR EMBL; X59964; CAA42590.1; -.
DR EMBL; X70377; CAA49838.1; -.
DR EMBL; AL591074; CAC94785.1; -.
DR PIR; A47142; A47142.
DR HSP; P01034; 1G96.
DR Genew; HGNC:2477; CSTS.
DR MIM; 123658; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.
DR InterPro; IPR00010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; Cy_1.
DR PROSITE; PS00287; CYSTATIN; 1.
DR Thiol protease inhibitor; Signal; Polymorphism.
KW SIGNAL.
FT CHAIN 1 20 PROBABLY.
FT ACT SITE 21 142 CYSTATIN D.
FT SITE 22 22 REACTIVE SITE (BY SIMILARITY).
FT DISULFID 95 105 SECONDARY AREA OF CONTACT.
FT DISULFID 119 139 BY SIMILARITY.
FT VARIANT 46 46 BY SIMILARITY.
FT C -> R (IN 45% OF THE POPULATION;
FT dbSNP:1799841).
FT /FTID=VAR 002208.
SQ SEQUENCE 142 AA; 16080 MM; CEFA89BA87A0DA68 CRC64;
Query Match 28.0%; Score 53; DB 1; Length 142;
Best Local Similarity 37.9%; Pred. No. 1.8;
Matches 11; Conservative 4; Mismatches 12; Indels 2; Gaps 1;
QY 1 QVTDHLEHNVEMQWTTCK--PETNC 27
DB 77 QIVGGVNYFNVKRGRTTCKSPNDNC 105
RESULT 15
MGR4_RAT STANDARD; PRT; 912 AA.
ID P31423;
AC 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Metabotropic glutamate receptor 4 precursor (MGLUR4).
GN GMR4 OR GPRC4 OR MGLUR4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1] SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
RX MEDLINE=92110002; Pubmed=1309649;
RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
RT "A family of metabotropic glutamate receptors.";
RL Neuron 11:41-52(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93338667; Pubmed=8338667;
RA O'Hara P.J., Sheppard P.O., Thøgersen H., Venezia D., Haldeman B.A.,
RA McGrane V., Houamed K.M., Thomsen C., Gilbert T.L., Mulvihill E.R.;
RT "The ligand-binding domain in metabotropic glutamate receptors is
RT related to bacterial periplasmic binding proteins.";
RL Neuron 11:41-52(1993).
RN [3]
RP INTERACTION WITH PRKCAB.
RX MEDLINE=20571397; Pubmed=11122333;
RA El Far O., Aitab J., Wischmeyer E., Nehring R.B., Karschin A.,
RA Betz H.;
RT "Interaction of the C-terminal tail region of the metabotropic
RT glutamate receptor 7 with the protein kinase C substrate PICK1.";
RL Eur. J. Neurosci. 12:4215-4221(2000).
CC -!- FUNCTION: Receptor for glutamate. The activity of this receptor
CC is mediated by a G-protein that inhibits adenylate cyclase
CC activity.

```

```

CC -!- SUBUNIT: Interacts with PRKCAB.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Is widely distributed in the CNS. Predominant
CC expression is seen in the granule cells of the cerebellum.
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR6.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M92077; -; NOT ANNOTATED_CDS.
DR EMBL; M90518; AA93150.1; -.
DR PIR; JH0563; JH0563.
DR InterPro; IPR001828; ANF_receptor.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS00979; G-PROTEIN_RECP_F3_1; 1.
DR PROSITE; PS00980; G-PROTEIN_RECP_F3_2; 1.
DR PROSITE; PS00981; G-PROTEIN_RECP_F3_3; 1.
DR PROSITE; PS50259; G-PROTEIN_RECP_F3_4; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 912 METABOTROPIC GLUTAMATE RECEPTOR 4.
FT DOMAIN 33 587 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 588 610 I (POTENTIAL).
FT DOMAIN 611 624 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 625 645 II (POTENTIAL).
FT DOMAIN 646 656 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 657 675 III (POTENTIAL).
FT DOMAIN 676 699 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 700 720 IV (POTENTIAL).
FT DOMAIN 721 750 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 751 772 V (POTENTIAL).
FT DOMAIN 773 785 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 786 808 VI (POTENTIAL).
FT DOMAIN 809 821 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 822 847 VII (POTENTIAL).
FT DOMAIN 848 912 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 484 484 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 124 124 Q -> R (IN REF. 2).
SQ SEQUENCE 912 AA; 101818 MM; 336430EF1B4B577 CRC64;
Query Match 28.0%; Score 53; DB 1; Length 912;
Best Local Similarity 39.0%; Pred. No. 12;
Matches 16; Conservative 3; Mismatches 6; Indels 16; Gaps 4;
QY 3 TDHLEHNVEMQWTTCK--PETNCVQGE 33
DB 496 TDHL-FLRIERMQPSSGOQLPRISICSLP---CQGERK 530

```

Search completed: March 18, 2004, 14:16:12  
Job time : 3.92857 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 19.7738 Seconds

(without alignments)  
526.560 Million cell updates/sec

Title: US-09-941-314-12

Perfect score: 189

Sequence: 1 QVTDHLEHYHNVEMQWTTCKPPTTNCVPOERE 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189	100.0	138	4	Q8WXU6
2	95	50.3	139	11	Q8KSA3
3	67	35.4	141	11	Q9DAP1
4	67	35.4	141	11	Q80ZM5
5	64	33.9	1779	5	Q18150
6	57	30.2	540	5	Q8T9G8
7	57	30.2	723	9	Q94M04
8	57	30.2	778	4	Q8TAA0
9	57	30.2	966	5	Q8T9K4
10	57	30.2	966	5	Q9VLM8
11	57	30.2	966	5	Q9USB4
12	56	29.6	140	11	Q9BXP9
13	55.5	29.4	633	5	Q9US49
14	54	28.6	315	10	Q9C7Z0
15	54	28.6	315	10	Q8LAV8
16	54	28.6	463	10	Q93Z38

17	54	28.6	463	10	Q9FX14	Q9FX14 arabidopsis
18	53.5	28.3	634	5	Q76188	Q76188 manduca sex
19	53	28.0	367	5	Q7YTM6	Q7YTM6 caenorhabdi
20	53	28.0	400	13	Q8UTR3	Q8UTR3 xenopus lae
21	53	28.0	983	11	Q62916	Q62916 rattus norv
22	52	27.5	3766	5	Q17551	Q17551 caenorhabdi
23	52	27.5	236	16	Q8DS42	Q8DS42 streptococc
24	52	27.5	441	5	Q962Y5	Q962Y5 aescaris suu
25	52	27.5	1347	12	Q8V7J8	Q8V7J8 chlorella v
26	51.5	27.2	367	10	Q9CGR6	Q9CGR6 arabidopsis
27	51.5	27.2	371	10	Q9C6F7	Q9C6F7 arabidopsis
28	51.5	27.2	678	2	Q83XK8	Q83XK8 legionella
29	51.5	27.2	1458	10	Q8W367	Q8W367 oryza sativ
30	51.5	27.2	1458	10	Q7XD96	Q7XD96 oryza sativ
31	51.5	27.2	2165	5	Q8T1P5	Q8T1P5 dictyosteli
32	51	27.0	270	16	Q7V796	Q7V796 prochloroco
33	51	27.0	392	5	Q9VYB3	Q9VYB3 dirosophila
34	51	27.0	591	16	Q8P9T0	Q8P9T0 leptospira
35	50.5	26.7	290	5	Q8SXB3	Q8SXB3 mus musculu
36	50.5	26.7	645	11	Q8C6W5	Q8C6W5 mus musculu
37	50	26.5	208	4	Q81YC9	Q81YC9 homo sapien
38	50	26.5	232	12	Q9YVN4	Q9YVN4 melanopis
39	50	26.5	261	5	Q17476	Q17476 hydraclonia
40	50	26.5	711	4	Q81ZM6	Q81ZM6 homo sapien
41	50	26.5	711	4	Q81U92	Q81U92 homo sapien
42	50	26.5	891	5	Q81B61	Q81B61 plasmodium
43	49.5	26.2	256	9	Q7Y4K3	Q7Y4K3 streptococc
44	49.5	26.2	622	13	Q91503	Q91503 torpeda mar
45	49.5	26.2	953	10	P93Z65	P93Z65 mesembryant

## ALIGNMENTS

## RESULT 1

ID	Q8WXU6	PRELIMINARY;	PRT;	138 AA.
AC	Q8WXU6	Q8WXU6		
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	SC13.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hail S.H.;			
RT	"SC13: A novel epididymal specific member of the cystatin family.";			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF335480; ALU7191.1; -			
DR	GO; GO:0004869; F:Cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR000010; Cystatin.			
DR	Pfam; PF00031; Cystatin; 1.			
DR	SMART; SM00043; Cy; 1			
SO	SEQUENCE	138 AA;	16506 MW;	E49440ACA3585C64 CRC64;

Query Match 100.0%; Score 189; DB 4; Length 138;  
Best Local Similarity 100.0%; Pred. No. 4.8e-20;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVTDHLEHYHNVEMQWTTCKPPTTNCVPOERE 33  
Db 76 QVTDHLEHYHNVEMQWTTCKPPTTNCVPOERE 108

## RESULT 2

ID Q8KSA3 PRELIMINARY; PRT; 139 AA.  
AC Q8KSA3;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)  
 DE Cystatin 11.  
 GN CSF11.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RA Hamil K.G., Hall S.H.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF501290; AAM21709.1; -  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 SQ SEQUENCE 139 AA; 1686 MW; E1E36DB786B4D08C CRC64;

Query Match 50.3%; Score 95; DB 11; Length 139;  
 Best Local Similarity 54.5%; Pred. No. 3.7e-06;  
 Matches 18; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 QVTDHLEHYLVNEMQWTTCK--PETTNCVPOE 33  
 ID Q9DAP1 PRELIMINARY; PRT; 141 AA.  
 DB 76 QMTNMFHITVEWQRTTCTKTEKMLCNVQGE 108

RESULT 3  
 Q9DAP1  
 ID Q9DAP1 PRELIMINARY; PRT; 141 AA.  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)  
 DE 1700006C19RIK protein.  
 GN 1700006C19RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schiraldi L.M., Scuderi F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fleischer C., Fujita M., Gariboldi M.F.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Marzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilting L.,  
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,  
 RA Hayashizaki Y.;  
 RL "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 DR EMBL; AK005665; BAB24175.1; -  
 DR HSSP; P01038; ICEW.  
 DR MGD; MGI:1916544; 1700006C19RIK.  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 SQ SEQUENCE 141 AA; 16811 MW; C20FA0DB8B1AC378C CRC64;

Query Match 35.4%; Score 67; DB 11; Length 141;  
 Best Local Similarity 42.4%; Pred. No. 0.051;  
 Matches 14; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 1 QVTDHLEHYLVNEMQWTTCK--PETTNCVPOE 31  
 ID Q80ZNS PRELIMINARY; PRT; 141 AA.  
 DB 76 QITDSLEYLVNEMQWTTCKKVGADNENCLFQ 108

RESULT 4  
 ID Q80ZNS PRELIMINARY; PRT; 141 AA.  
 AC Q80ZNS;  
 DT 01-JUN-2003 (TReMBLrel. 24, Created)  
 DT 01-JUN-2003 (TReMBLrel. 24, last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)  
 DE RIKEN cDNA 1700006C19 gene.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testicle;  
 RA Strauberg R.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC048681; AAH48681.1; -  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR InterPro; IPR003243; Cystatin\_C/M.  
 DR Pfam; PF00031; Cystatin; 1.  
 DR ProDom; PD001231; Cystatin\_C/M; 1.  
 DR SMART; SM00043; CY; 1.  
 SQ SEQUENCE 141 AA; 16825 MW; C20FA0DB8B484951F CRC64;

Query Match 35.4%; Score 67; DB 11; Length 141;  
 Best Local Similarity 42.4%; Pred. No. 0.051;  
 Matches 14; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 1 QVTDHLEHYLVNEMQWTTCK--PETTNCVPOE 31  
 ID Q18150 PRELIMINARY; PRT; 1779 AA.  
 DB 76 QITDSLEYLVNEMQWTTCKKVGADNENCLFQ 108

RESULT 5  
 ID Q18150 PRELIMINARY; PRT; 1779 AA.  
 AC Q18150; Q45627;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 08, last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)  
 DE T28B8.4 protein.  
 GN T28B8.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA White S.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkin T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirtlen J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sultson J.,

RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA White S.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z81133; CAB03445.1; -;  
 DR EMBL; AL021066; CAB03445.1; JOINED.  
 DR EMBL; AL021066; CA15925.1; -;  
 DR EMBL; Z81133; CA15925.1; JOINED.  
 DR PIR; T23130; T23130.  
 DR Wormpep; T2888.4; CE16519.  
 DR Interpro; IPR008938; ANM.  
 SQ SEQUENCE 1779 AA; 205511 MW; 70A4489148B96F3E CRC64;

Query Match 33.9%; Score 64; DB 5; Length 1779;  
 Best Local Similarity 34.2%; Pred. No. 1.9;  
 Matches 13; Conservative 8; Mismatches 7; Indels 10; Gaps 2;

Qy 1 QVTDHLEVHLN-----VEMQWTCCKPPTTCVPO 30  
 Db 750 RITDHLDIINNLPTLKHPTNVQWIT--DPTPHCLPK 785

RESULT 6  
 Q8T9G8 PRELIMINARY; PRT; 540 AA.  
 AC Q8T9G8;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE S001519P.  
 GN AATS-ALA OR CG13391.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 OX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stapleton M., Broketein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY069773; AAL39918.1; -;  
 DR FLYBase; FBgn0027094; Aats-ala.  
 DR GO; GO:0004813; F:alanine-tryptophan ligase activity; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0006419; P:amyl-tryptophan aminocyclization; IEA.  
 DR Interpro; IPR003156; DHHA1.  
 DR Interpro; IPR002318; CRNA-synt-2c.  
 DR Interpro; IPR006193; CRNA-synt-ala.  
 DR Pfam; PF02272; DHHA1; 1.  
 DR Pfam; PF01411; CRNA-synt-2c; 1.  
 DR PROSITE; PS50860; AA-TRNA\_LIGASE\_II\_ALA; 1.  
 SQ SEQUENCE 540 AA; 59580 MW; B870DF9C386843B CRC64;

Query Match 30.2%; Score 57; DB 5; Length 540;  
 Best Local Similarity 35.7%; Pred. No. 6;  
 Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy 1 QVTDHLEVHLNEMQWTCCKPPTTCVPO 28  
 Db 157 KVGDELHLDIVERRMLTKNHSATHAL 184

RESULT 7

Q94M04  
 ID Q94M04 PRELIMINARY; PRT; 723 AA.  
 AC Q94M04;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Major core structural protein Pl.  
 GN 1.  
 OS Bacteriophage phi-12.  
 OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.  
 OX NCBI\_Taxid=161736;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gottlieb P.J., Potgieter C., Wei H., Toporovsky I.;  
 RT "Characterization of Bacteriophage phi12.";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF408636; AAL01109.1; -;  
 DR Interpro; IPR006162; Phantme\_S.  
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.  
 SQ SEQUENCE 723 AA; 78950 MW; F5DEF7FF93A50B7C CRC64;

Query Match 30.2%; Score 57; DB 9; Length 723;  
 Best Local Similarity 50.0%; Pred. No. 8.1;  
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 5 HLEVHLNEMQWTCCKP 22  
 Db 463 HTEYRFRVDLQWATCYLP 480

RESULT 8  
 Q8TA0 PRELIMINARY; PRT; 778 AA.  
 ID Q8TA0;  
 AC Q8TA0;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Similar to PRAM-1 protein, PML-RARA target gene encoding an adaptor  
 DE molecule-1.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tissue-Testis;  
 RA Strauberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC026171; AAL26171.1; -;  
 SQ SEQUENCE 778 AA; 89091 MW; 405B01953B2704DD CRC64;

Query Match 30.2%; Score 57; DB 4; Length 778;  
 Best Local Similarity 38.7%; Pred. No. 8.7;  
 Matches 12; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

Qy 1 QVTDHLEVHLNEMQWTCCKPPTTCVPO 31  
 Db 148 QVTLGHLPERLEDAWACETQERTTEVPT 178

RESULT 9  
 Q8T9K4 PRELIMINARY; PRT; 966 AA.  
 ID Q8T9K4;  
 AC Q8T9K4;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE GMO3058P.  
 GN AATS-ALA OR CG13391.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.

```

OX NCBI_TaxID=7227;
RN
  [1]
  SEQUENCE FROM N.A.
  RA Stapleton M., Brockstein P., Hong L., Agbayan A., Carlson J.,
  RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
  RA Gonzalez M., Guerin H., Li P., Liao G., Miranda A., Mungall C.J.,
  RA Nuno J., Paclet J., Paragas V., Park S., Phouanavong S., Wan K.,
  RA Yu C., Lewis S.E., Rubin G.M., Celinker S.
  RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
  DR EMBL: AY069255; ALU39400.1; -
  DR FLYBase: FBgn0027094; Aats-ala.
  DR GO: GO:0004813; F:alanine-tRNA ligase activity; IEA.
  DR GO: GO:0005524; F:ATP binding; IEA.
  DR GO: GO:0003676; F:nucleic acid binding; IEA.
  DR GO: GO:0006419; P:alanyl-tRNA aminoacylation; IEA.
  DR InterPro: IPR003156; DHHA1.
  DR InterPro: IPR006193; tRNA-synt_2c.
  DR InterPro: IPR006193; tRNA-synt_ala.
  DR Pfam: PF02272; DHHA1; 1.
  DR Pfam: PF01411; tRNA-synt_2c; 1.
  DR PRINTS: PRO0980; TRNASYNTHALA.
  DR TIGRFAMs: TIGR00344; alas; 1.
  DR PROSITE: PSS0860; AA TRNA LIGASE II ALA; 1.
SQ
  SEQUENCE 966 AA; 107697 MW; 9F2079A119F463B4 CRC64;

Query Match      30.2%; Score 57; DB 5; Length 966;
Best Local Similarity 35.7%; Pred. No. 11;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCCQPERTTCV 28
DB 583 KVGDLELHIDVERRWLTMKHNSATHAL 610

Q9U6B4
ID Q9U6B4 PRELIMINARY; PRT; 966 AA.
AC Q9U6B4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AATS-ALA protein.
GN AATS-ALA OR CG13391.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN
  [1]
  SEQUENCE FROM N.A.
  RC STRAIN=Berkely;
  RX MEDLINE=20196006; PubMed=10731132;
  RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
  RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
  RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
  RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
  RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
  RA Man K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
  RA Abail J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
  RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
  RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov E.M.,
  RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
  RA Burdick K.C., Butan D.A., Butler H., Cadieu E., Center A., Chandra I.,
  RA Cherry J.M., Cleyer S., Dahlke C., Davenport L.B., Davies P.,
  RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
  RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
  RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
  RA Fouts L., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
  RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
  RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
  RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwan C.,
  RA Jaitani M., Kalish C., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
  RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

```

  RA Lascko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
  RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
  RA Merkulov G., Milshina N.V., Mobarry B., Morris J., Moshrefi A.,
  RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
  RA Nelson S.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
  RA Palazzolo M., Peltman G.S., Fan S., Pollard J., Furl V., Reese M.G.,
  RA Rehnert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
  RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
  RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
  RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
  RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
  RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
  RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
  RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
  RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
  RT "The genome sequence of Drosophila melanogaster."
  RL Science 287:2185-2195(2000).
  DR EMBL: AE003621; AAF52657.1; -
  DR FLYBase: FBgn0027094; Aats-ala.
  DR GO: GO:0004813; F:alanine-tRNA ligase activity; IEA.
  DR GO: GO:0005524; F:ATP binding; IEA.
  DR GO: GO:0003676; F:nucleic acid binding; IEA.
  DR GO: GO:0006419; P:alanyl-tRNA aminoacylation; IEA.
  DR InterPro: IPR003156; DHHA1.
  DR InterPro: IPR003156; tRNA-synt_2c.
  DR InterPro: IPR002318; tRNA-synt_ala.
  DR InterPro: IPR006193; tRNA-synt_ala.
  DR Pfam: PF02272; DHHA1; 1.
  DR Pfam: PF01411; tRNA-synt_2c; 1.
  DR PRINTS: PRO0980; TRNASYNTHALA.
  DR TIGRFAMs: TIGR00344; alas; 1.
  DR PROSITE: PSS0860; AA TRNA LIGASE II ALA; 1.
SQ
  SEQUENCE 966 AA; 107741 MW; B1B63A7A8B87B6E CRC64;

Query Match      30.2%; Score 57; DB 5; Length 966;
Best Local Similarity 35.7%; Pred. No. 11;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCCQPERTTCV 28
DB 583 KVGDLELHIDVERRWLTMKHNSATHAL 610

Q9U6B4
ID Q9U6B4 PRELIMINARY; PRT; 966 AA.
AC Q9U6B4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alanyl-tRNA synthetase.
GN AATS-ALA OR ALAS OR CG13391.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN
  [1]
  SEQUENCE FROM N.A.
  RP Chihade J.W., Brown J.R., Schimmel P., Ribas de Pouplana L.,
  RT "Detection of an intermediate stage of Mitochondria Genesis."
  RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
  DR EMBL: AF188718; AAF05593.1; -
  DR FLYBase: FBgn0027094; Aats-ala.
  DR GO: GO:0004813; F:alanine-tRNA ligase activity; IEA.
  DR GO: GO:0005524; F:ATP binding; IEA.
  DR GO: GO:0003676; F:nucleic acid binding; IEA.
  DR GO: GO:0006419; P:alanyl-tRNA aminoacylation; IEA.
  DR InterPro: IPR003156; DHHA1.
  DR InterPro: IPR002318; tRNA-synt_2c.
  DR InterPro: IPR006193; tRNA-synt_ala.
  DR Pfam: PF02272; DHHA1; 1.
  DR PRINTS: PRO0980; TRNASYNTHALA.

```

DR TIGRFBMS; TIGR00344; alas; 1.  
 DR PROSITE; PSS0860; AA\_TRNA\_LIGASE\_II\_ALA; 1.  
 KM Antisense; tRNA synthetase.  
 SQ SEQUENCE 966 AA; 107877 MW; 82034A189F0C81B5 CRC64;

Query Match 30.2%; Score 57; DB 5; Length 966;  
 Best Local Similarity 35.7%; Pred. No. 11;  
 Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 QVTDHLRYHLNEMQWTTCKPFTNCV 28  
 Db 583 KVGDELEHIDVERHMTMKNSATHAL 610

RESULT 12

Q9EPX9 PRELIMINARY; PRT; 140 AA.  
 AC Q9EPX9; 01-MAR-2001 (TIGRFBMS; 16, Created)  
 DT 01-MAR-2001 (TIGRFBMS; 16, Last sequence update)  
 DT 01-JUN-2003 (TIGRFBMS; 24, Last annotation update)  
 DE Cystatin C.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C;  
 RX MEDLINE=21010502; PubMed=11144350;  
 RA Taupin P.J., Ray J., Fischer W.H., Suh S.T., Hakansson K., Grubb A.,  
 RT "RFG-2-Responsive neural stem cell proliferation requires CCG, a novel  
 RT autocorin/paracrine cofactor.";  
 RT Neuron 28:385-397(2000).  
 DR EMBL; AF311741; AAC40283.1; -  
 DR HSSP; P01034; 1G96.  
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 FT CHAIN 21 140  
 FT VARIANT 16 A -> G.  
 FT VARIANT 84 L -> F.  
 FT VARIANT 84  
 SQ SEQUENCE 140 AA; 15517 MW; 3A563406D58D785 CRC64;

Query Match 29.6%; Score 56; DB 11; Length 140;  
 Best Local Similarity 44.8%; Pred. No. 2.1;  
 Matches 13; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

QY 1 QVTDHLRYHLNEMQWTTCKPFTNCV 27  
 Db 75 QLVAGVNYFLDVEWGRTTKSGTNLTDC 103

RESULT 13

Q9U5A9 PRELIMINARY; PRT; 633 AA.  
 AC Q9U5A9; 01-MAY-2000 (TIGRFBMS; 13, Created)  
 DT 01-MAY-2000 (TIGRFBMS; 13, Last sequence update)  
 DT 01-JUN-2003 (TIGRFBMS; 24, Last annotation update)  
 DE Amino acid transporter/amino acid-gated channel for sodium/potassium  
 DE ions.  
 GN CAATCH1.  
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Dictyssa; Sphingioidea;  
 OC Sphingidae; Sphinginae; Manduca.  
 NCBI\_Taxid=7130;  
 [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Midgut;  
 RX MEDLINE=20387341; PubMed=10829035;  
 RA Feldman D.H., Harvey W.R., Stevens B.R.;  
 RT "A novel electrogenic amino acid transporter is activated by K+ or  
 RT Na+, is alkaline pH-dependent, and is Cl--independent.";  
 RL J. Biol. Chem. 275:24518-24526(2000).  
 DR EMBL; AF013963; AAF18560.1; -  
 DR GO; GO:0005087; C:integral to plasma membrane; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005328; F:neurotransmitter:sodium symporter activity; IEA.  
 DR GO; GO:0006836; F:neurotransmitter transport; IEA.  
 DR InterPro; IPR000175; Na/ntra\_n\_symport.  
 DR Pfam; PF00209; SNF; 1.  
 DR PRINTS; PR00176; NANEUSMPORT.  
 DR ProDom; PD000448; Na/ntra\_n\_symport; 1.  
 DR PROSITE; PSS0267; NA\_NEUROTRAN\_SYM\_3; 1.  
 SQ SEQUENCE 633 AA; 69934 MW; 597EDC58E98FDPID CRC64;

Query Match 29.4%; Score 55.5; DB 5; Length 633;  
 Best Local Similarity 42.9%; Pred. No. 12;  
 Matches 12; Conservative 3; Mismatches 8; Indels 5; Gaps 2;

QY 8 YHLNEMQWTTCKPFTNCVPOB 31  
 Db 156 YILMSFQATLPNALICQ-PEMENVCPSD 182

RESULT 14

Q9C7Z0 PRELIMINARY; PRT; 302 AA.  
 AC Q9C7Z0;  
 DT 01-JUN-2001 (TIGRFBMS; 17, Created)  
 DT 01-JUN-2001 (TIGRFBMS; 17, Last sequence update)  
 DT 01-JUN-2003 (TIGRFBMS; 24, Last annotation update)  
 DE Desiccation-related protein, putative.  
 GN T2J15.11.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_Taxid=3702;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,  
 RA Dunn P., Eggu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gali J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
 RA Millechner J., Miranda M., Nguyen M., Nierman W.C., Osborne B.L.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utechtack T., Van Aken S., Vaynsberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Frazer C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:816-820(2000).  
 DR EMBL; AC051631; AAG51530.1; -  
 DR PIR; B96520; B96520.  
 SQ SEQUENCE 302 AA; 32808 MW; ED2B80339ECB64E4 CRC64;

Query Match 28.6%; Score 54; DB 10; Length 302;  
 Best Local Similarity 42.3%; Pred. No. 9.2;  
 Matches 11; Conservative 4; Mismatches 7; Indels 4; Gaps 1;

QY 8 YHLNEMQWTTCKPFTNCVPOB 33

DB 11 HHLQV---TSCPDQATTNCTDQDRK 32

## RESULT 15

Q8LAU8 PRELIMINARY; PRT; 315 AA.  
 AC Q8LAU8;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Desiccation-related protein, putative.  
 OS Arabidopsis thaliana (mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Haas B.U., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 RT annotation.";  
 RL Genome Biol. 0:0-0(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RT "Full-length cDNA from Arabidopsis thaliana."  
 RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AY087598; AM65140.1; -;  
 SQ SEQUENCE 315 AA; 34328 MW; 5AFPCA9D6391F715 CRC64;

Query Match 28.6%; Score 54; DB 10; Length 315;  
 Best local Similarity 42.3%; Pred. No. 9.6;  
 Matches 11; Conservative 4; Mismatches 7; Indels 4; Gaps 1;

QY 8 YHLNVEQMWTTCQKPEFTNCVQPERE 33  
 DB 24 HHLQV---TSCPDQATTNCTDQDRK 45

Search completed: March 18, 2004, 14:21:31  
 Job time : 20.7738 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 44.4306 Seconds  
(without alignments)  
311.606 Million cell updates/sec

Title: US-09-941-314-13  
Perfect score: 272  
Sequence: 1 DQYKXSDDKYHFRIFRVLK.....EYHLNVMQWTCCKPPTN 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	272	100.0	49	5 AAU79863	AAU79863 Human cys
2	272	100.0	115	5 AAU79853	AAU79853 Human cys
3	272	100.0	117	5 AAU79854	AAU79854 Human cys
4	272	100.0	137	5 AAU79852	AAU79852 Human cys
5	266	97.8	80	5 AAU79865	AAU79865 Human cys
6	254	93.4	46	5 AAU79860	AAU79860 Human cys
7	248	91.2	52	5 AAU79864	AAU79864 Human cys
8	154	56.6	59	5 AAU79866	AAU79866 Human cys
9	149	54.8	33	5 AAU79862	AAU79862 Human cys
10	143	52.6	27	5 AAU79859	AAU79859 Human cys
11	143	52.6	35	5 AAU79858	AAU79858 Human cys
12	138	50.7	24	5 AAU79861	AAU79861 Human cys
13	128	47.1	50	4 AAM15096	AAM15096 Peptide #
14	128	47.1	50	4 AAB34086	AAB34086 Peptide #
15	128	47.1	50	4 AAM27545	AAM27545 Peptide #
16	128	47.1	50	4 AAB33289	AAB33289 Peptide #
17	128	47.1	50	4 ABB28913	ABB28913 Peptide #
18	128	47.1	50	4 ABB19524	ABB19524 Protein #
19	128	47.1	50	4 AAM67252	AAM67252 Human bon
20	128	47.1	50	4 AAM54871	AAM54871 Human bra
21	128	47.1	50	4 ABB48915	ABB48915 Human liv
22	128	47.1	50	4 AAM02833	AAM02833 Peptide #
23	128	47.1	50	5 ABB36903	ABB36903 Human pep
24	127	46.7	92	2 AAW78259	AAW78259 Fragment
25	127	46.7	113	6 ADA57563	ADA57563 Human sec

## ALIGNMENTS

26	127	46.7	113	6 ADA41457	ADA41457 Human sec
27	127	46.7	113	7 ADC74577	ADC74577 Human sec
28	127	46.7	113	7 ADD38088	ADD38088 Human sec
29	127	46.7	114	2 AAW78153	AAW78153 Human sec
30	127	46.7	123	2 AAW78260	AAW78260 Fragment
31	127	46.7	142	2 AAW78258	AAW78258 Human cys
32	127	46.7	142	4 AAB02405	AAB02405 Human cys
33	127	46.7	142	4 AAB04434	AAB04434 Human cys
34	127	46.7	142	6 ADA57231	ADA57231 Human sec
35	127	46.7	142	6 ADA41112	ADA41112 Human sec
36	127	46.7	142	7 ADC74335	ADC74335 Human sec
37	127	46.7	142	7 ADD37980	ADD37980 Human pro
38	127	46.7	142	7 ADD46706	ADD46706 Human pro
39	127	46.7	142	7 ADD46710	ADD46710 Human pro
40	122	44.9	141	3 AAY96576	AAY96576 Murine cy
41	122	44.9	141	4 AAB02403	AAB02403 Murine cy
42	122	44.9	141	4 AAB04432	AAB04432 Mouse tes
43	120	44.1	142	4 AAB02404	AAB02404 Murine cy
44	120	44.1	142	4 AAB04433	AAB04433 Mouse cys
45	120	44.1	143	6 ADA14374	ADA14374 Mouse spe

## RESULT 1

AAU79863 standard; peptide; 49 AA.

AAU79863;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcy8) antigenic fragment #11.

Cystatin-8; Zcy8; cancer; procoagulant protein; thrombosis;

spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

MO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001MO-US026668.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.

Claim 2; Page 97-98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcy8) polypeptide (I). (I) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcy8 is useful as research reagent for characterising sites of interaction between Zcy8 and its receptor. Zcy8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(I) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcy8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (1) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (1) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
SQ Sequence 49 AA;  
Query Match 100.0%; Score 272; DB 5; Length 49;  
Best Local Similarity 100.0%; Pred. No. 4e-30;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DOWNKESDDKXHFRIFFVLKVRQVTDHLEHNLVEMQMTTCKQPEPTN 49  
Db 1 DOWNKESDDKXHFRIFFVLKVRQVTDHLEHNLVEMQMTTCKQPEPTN 49  
RESULT 2  
ID AAU79853 standard; protein; 115 AA.  
XX  
AC AAU79853;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human cystatin-8 (Zcys8) antigenic fragment #1.  
XX  
KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KM sperm motility; fertilisation; antigenic fragment.  
XX  
OS Homo sapiens.  
XX  
PN WO200220567-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US026868.  
XX  
PR 01-SEP-2000; 2000US-0230230P.  
XX  
PA (ZYMO) ZYMOGENETICS INC.  
XX  
PI Holloway JL, Gao Z, Bishop PD;  
XX  
XX WPI; 2002-383044/41.  
XX  
DR WPI; 2002-383044/41.  
XX  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX  
PS Claim 2; Page 94; 100pp; English.  
XX  
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(1) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (1) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (1) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic fragment of human cystatin-8  
CC (Zcys8)  
XX  
SQ Sequence 115 AA;  
Query Match 100.0%; Score 272; DB 5; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1.1e-29;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DOWNKESDDKXHFRIFFVLKVRQVTDHLEHNLVEMQMTTCKQPEPTN 49  
Db 30 DOWNKESDDKXHFRIFFVLKVRQVTDHLEHNLVEMQMTTCKQPEPTN 78  
RESULT 3  
ID AAU79854 standard; protein; 117 AA.  
XX  
AC AAU79854;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human cystatin-8 (Zcys8) antigenic fragment #2.  
XX  
KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KM sperm motility; fertilisation; antigenic fragment.  
XX  
OS Homo sapiens.  
XX  
PN WO200220567-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US026868.  
XX  
PR 01-SEP-2000; 2000US-0230230P.  
XX  
PA (ZYMO) ZYMOGENETICS INC.  
XX  
PI Holloway JL, Gao Z, Bishop PD;  
XX  
XX WPI; 2002-383044/41.  
XX  
DR WPI; 2002-383044/41.  
XX  
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.  
XX  
PS Claim 2; Page 94-95; 100pp; English.  
XX  
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis, modulating seminal  
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(1) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (1) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (1) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and

CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic fragment of human cystatin-8  
 CC (Zcys8)  
 XX Sequence 117 AA;  
 SQ Query Match 100.0%; Score 272; DB 5; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-29;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DQYKESDDKXHPRIFFVLKVRQVTDHLEHYHNVEMQWTTCKRPETTN 49  
 32 DQYKESDDKXHPRIFFVLKVRQVTDHLEHYHNVEMQWTTCKRPETTN 80  
 Db  
 RESULT 4  
 AAU79852  
 ID AAU79852 standard; protein; 137 AA.  
 AC AAU79852;  
 XX  
 XX 15-JUL-2002 (first entry)  
 DT  
 XX Human cystatin-8 (Zcys8).  
 DE  
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation.  
 OS Homo sapiens.  
 XX  
 XX MO200220567-A2.  
 PN  
 XX 14-MAR-2002.  
 PD  
 XX 29-AUG-2001; 2001MO-US026868.  
 PF  
 XX 01-SEP-2000; 2000US-0230230P.  
 PR  
 XX (ZYMO ) ZYMOGENETICS INC.  
 PA  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 XX WPI: 2002-383044/41.  
 DR N-PSDB; ABK49522.  
 XX  
 XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 PS Claim 2; Page 93-94; 100pp; English.  
 XX  
 XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 gene.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)

XX Sequence 137 AA;  
 SQ Query Match 100.0%; Score 272; DB 5; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-29;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DQYKESDDKXHPRIFFVLKVRQVTDHLEHYHNVEMQWTTCKRPETTN 49  
 52 DQYKESDDKXHPRIFFVLKVRQVTDHLEHYHNVEMQWTTCKRPETTN 100  
 Db  
 RESULT 5  
 AAU79865  
 ID AAU79865 standard; peptide; 80 AA.  
 AC AAU79865;  
 XX  
 XX 15-JUL-2002 (first entry)  
 DT  
 XX Human cystatin-8 (Zcys8) antigenic fragment #13.  
 DE  
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.  
 OS Homo sapiens.  
 XX  
 XX MO200220567-A2.  
 PN  
 XX 14-MAR-2002.  
 PD  
 XX 29-AUG-2001; 2001MO-US026868.  
 PF  
 XX 01-SEP-2000; 2000US-0230230P.  
 PR  
 XX (ZYMO ) ZYMOGENETICS INC.  
 PA  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 XX WPI: 2002-383044/41.  
 DR  
 XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 PS Claim 2; Page 98; 100pp; English.  
 XX  
 XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 gene.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 XX Sequence 80 AA;  
 SQ Query Match 97.8%; Score 266; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 4.9e-29;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QVNESDCKYHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 49  
DB 1 QVNESDCKYHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 48

## RESULT 6

AAU79860  
ID AAU79860 standard; peptide: 46 AA.

XX AAU79860;

XX 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcys8) antigenic fragment #8.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX MO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001MO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 97; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis; modulating seminal  
CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
XX Sequence 46 AA;

QY Query Match 93.4%; Score 254; DB 5; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.2e-27;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NKESDCKYHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 49

DB 1 NKESDCKYHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 46

## RESULT 7

AAU79864  
ID AAU79864 standard; peptide: 52 AA.

XX AAU79864;

XX 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcys8) antigenic fragment #12.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX MO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001MO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 98; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)  
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
CC protein in an individual and thus inhibiting the thrombotic events  
CC associated with cancer; promoting spermatogenesis; modulating seminal  
CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm  
CC motility and fertilisation; and as antigenic peptides to generate  
CC antibodies. Zcys8 is useful as research reagent for characterising sites  
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
CC enhancing fertilisation during assisted reproduction in humans and in  
CC animals. Anti-(I) antibodies are useful to screen biological samples like  
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
CC presence of Zcys8. The antibodies are also useful to isolate large  
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
CC The polynucleotide encoding (I) is useful to detect and to localise the  
CC expression of a Zcys8 gene in a biological sample and Zcys8  
CC oligonucleotide probes are useful for in vivo diagnosis. The  
CC polynucleotide encoding (I) is useful in determining whether a subject's  
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
CC copy number changes, insertions, deletions, restriction site changes and  
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
XX  
XX Sequence 52 AA;

QY Query Match 91.2%; Score 248; DB 5; Length 52;  
Best Local Similarity 100.0%; Pred. No. 9.2e-27;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KESDCKYHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 49  
DB 1 KESDCKYHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 45

RESULT 8  
 ID AAV79866 standard; peptide; 59 AA.  
 AC AAV79866;  
 XX  
 AC AAV79866;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #14.  
 XX  
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 99; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 59 AA;  
 XX  
 Query Match 56.6%; Score 154; DB 5; Length 59;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-13;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 23 QVTDHLEHYHLNVEMQWTTCKRPETTN 49  
 Db 1 QVTDHLEHYHLNVEMQWTTCKRPETTN 27  
 XX  
 RESULT 9  
 ID AAV79862 standard; peptide; 33 AA.  
 XX

AC AAV79862;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human cystatin-8 (Zcys8) antigenic fragment #10.  
 XX  
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;  
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;  
 KM sperm motility; fertilisation; antigenic peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX  
 DR WPI; 2002-383044/41.  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 97; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis, modulating seminal  
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterising sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 XX  
 SQ Sequence 33 AA;  
 XX  
 Query Match 54.8%; Score 149; DB 5; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-13;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 24 QVTDHLEHYHLNVEMQWTTCKRPETTN 49  
 Db 1 QVTDHLEHYHLNVEMQWTTCKRPETTN 26  
 XX  
 RESULT 10  
 ID AAV79859 standard; peptide; 27 AA.  
 AC AAV79859;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX



PN WO200220567-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 29-AUG-2001; 2001WO-US026868.  
 XX  
 PR 01-SEP-2000; 2000US-0230230P.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Holloway JL, Gao Z, Bishop PD;  
 XX WPI; 2002-383044/41.  
 DR  
 XX  
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting  
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads  
 PT to inhibition of thrombotic events associated with cancer.  
 XX  
 PS Claim 2; Page 97; 100pp; English.  
 XX  
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)  
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant  
 CC protein in an individual and thus inhibiting the thrombotic events  
 CC associated with cancer; promoting spermatogenesis; modulating seminal  
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm  
 CC motility and fertilisation; and as antigenic peptides to generate  
 CC antibodies. Zcys8 is useful as research reagent for characterizing sites  
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in  
 CC enhancing fertilisation during assisted reproduction in humans and in  
 CC animals. Anti-(I) antibodies are useful to screen biological samples like  
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the  
 CC presence of Zcys8. The antibodies are also useful to isolate large  
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.  
 CC The polynucleotide encoding (I) is useful to detect and to localise the  
 CC expression of a Zcys8 gene in a biological sample and Zcys8  
 CC oligonucleotide probes are useful for in vivo diagnosis. The  
 CC polynucleotide encoding (I) is useful in determining whether a subject's  
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene  
 CC copy number changes, insertions, deletions, restriction site changes and  
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.  
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)  
 CC  
 SO Sequence 24 AA;  
 Query Match 50.7%; Score 138; DB 5; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-12;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 23 ROYTDHLEYHLNVMQWTTCKPE 46  
 DB 1 ROYTDHLEYHLNVMQWTTCKPE 24  
 RESULT 13  
 AAM15096  
 ID AAM15096 standard; protein; 50 AA.  
 AC AAM15096;  
 XX  
 XX 12-OCT-2001 (first entry)  
 DT  
 XX  
 DE Peptide #1530 encoded by probe for measuring cervical gene expression.  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 XX cervical cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000670.  
 XX

XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488901/53.  
 DR  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human cervical epithelial cells.  
 XX  
 PS Claim 27; SEQ ID NO 19922; 487pp; English.  
 XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENPs; see A4110068-A4128459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix; notably  
 CC cervical cancer. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 SO Sequence 50 AA;  
 Query Match 47.1%; Score 128; DB 4; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-10;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DQYKESDDYKHFRIPLVVKORO 24  
 DB 27 DQYKESDDYKHFRIPLVVKORO 50  
 RESULT 14  
 ABB34086  
 ID ABB34086 standard; peptide; 50 AA.  
 AC ABB34086;  
 XX  
 XX 04-FEB-2002 (first entry)  
 DT  
 XX  
 DE Peptide #1592 encoded by human foetal liver single exon probe.  
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000669.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX

XX WP1; 2001-483447/52.  
DR  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human fetal liver.  
XX  
PS Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for measuring  
CC human gene expression in a sample derived from human foetal liver. The  
CC single exon nucleic acid probes may be used for predicting, measuring and  
CC displaying gene expression in samples derived from human fetal liver. The  
CC present sequence is a peptide encoded by a single exon nucleic acid probe  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)  
XX  
SQ Sequence 50 AA;  
  
Query Match 47.1%; Score 128; DB 4; Length 50;  
Best Local Similarity 100.0%; Fred.No. 3.9e-10;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
  
Oy 1 DOYNKESDDKYHFRIFRVLKVORQ 24  
|||  
Db 27 DOYNKESDDKYHFRIFRVLKVORQ 50  
  
RESULT 15  
AAM27545  
ID AAM27545 standard; protein; 50 AA.  
XX  
AC AAM27545;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #1582 encoded by probe for measuring placental gene expression.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder.  
XX  
OS Homo sapiens.  
XX  
FN WO200157272-A2.  
PN  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000663.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WP1; 2001-488897/53.  
PT  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human placenta.  
XX  
PS Claim 27; SEQ ID NO 27814; 654pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENP:  
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of

```
CC      human genetic disorders
XX
SQ      Sequence 50 AA;
        Query Match          47.1%; Score 128; DB 4; Length 50;
        Best Local Similarity 100.0%; Pired. No. 3.9e-10;
        Matches    24; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy      1 DQYNKESDDKKYHRIIFRLVKYQRQ 24
         |||||
Db      . 27 DQYNKESDDKKYHRIIFRLVKYQRQ 50
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 12.5417 Seconds  
(Without alignments)  
201.701 Million cell updates/sec

Title: US-09-941-314-13

Perfect score: 272  
Sequence: 1 DOWNKESDCKYHFRFVLK.....EHLVNMOWTTCOKPETT 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/2/1aa/5A COMB pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B COMB pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A COMB pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B COMB pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS COMB pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1 pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	46.7	142	3	US-09-431-480-4 Sequence 4, Appl
2	127	46.7	142	3	US-09-617-302-4 Sequence 4, Appl
3	122	44.9	141	3	US-09-431-480-2 Sequence 2, Appl
4	122	44.9	141	3	US-09-617-302-2 Sequence 2, Appl
5	120	44.1	142	3	US-09-431-480-3 Sequence 3, Appl
6	120	44.1	142	3	US-09-617-302-3 Sequence 3, Appl
7	112	41.2	127	4	US-08-849-303-19 Sequence 13, Appl
8	108	39.7	140	4	US-09-886-319A-46 Sequence 46, Appl
9	108	39.7	140	4	US-09-886-319A-48 Sequence 48, Appl
10	104	38.2	140	3	US-09-431-480-5 Sequence 5, Appl
11	104	38.2	140	3	US-09-617-302-5 Sequence 5, Appl
12	104	38.2	140	4	US-08-849-303-18 Sequence 18, Appl
13	103	37.9	112	4	US-08-849-303-16 Sequence 16, Appl
14	103	37.9	116	4	US-09-775-932-16 Sequence 16, Appl
15	103	37.9	118	4	US-09-775-932-24 Sequence 24, Appl
16	103	37.9	120	4	US-09-775-932-2 Sequence 2, Appl
17	103	37.9	120	6	US-08-791-522-4 Patent No. 5432264
18	103	37.9	139	2	US-08-791-522-4 Sequence 4, Appl
19	103	37.9	139	4	US-09-314-777-4 Sequence 4, Appl
20	103	37.9	139	4	US-08-849-303-15 Sequence 15, Appl
21	103	37.9	145	2	US-08-832-535-11 Sequence 11, Appl
22	103	37.9	146	2	US-08-791-522-3 Sequence 3, Appl
23	103	37.9	146	2	US-08-744-138-3 Sequence 3, Appl
24	103	37.9	146	3	US-09-019-485-4 Sequence 4, Appl
25	103	37.9	146	3	US-09-314-777-3 Sequence 3, Appl
26	103	37.9	146	3	US-09-431-480-6 Sequence 6, Appl
27	103	37.9	146	3	US-09-617-302-6 Sequence 6, Appl

28	103	37.9	146	4	US-09-241-376-3 Sequence 3, Appl
29	103	37.9	146	4	US-09-528-436B-3 Sequence 3, Appl
30	103	37.9	146	4	US-09-886-319A-47 Sequence 47, Appl
31	103	37.9	146	4	US-09-940-497-3 Sequence 3, Appl
32	103	37.9	146	4	US-09-976-594-37 Sequence 37, Appl
33	103	37.9	146	4	US-08-849-303-17 Sequence 17, Appl
34	103	37.9	146	5	PCT-US95-07135-9 Sequence 9, Appl
35	103	37.9	146	6	5432264-6 Patent No. 5432264
36	93	34.2	121	4	US-09-775-932-8 Sequence 8, Appl
37	93	34.2	111	3	US-08-744-138-6 Sequence 6, Appl
38	93	34.2	141	4	US-09-241-376-6 Sequence 6, Appl
39	93	34.2	141	4	US-09-940-497-6 Sequence 6, Appl
40	93	34.2	141	4	US-08-849-303-24 Sequence 24, Appl
41	90	33.1	111	4	US-08-849-303-26 Sequence 26, Appl
42	89	32.7	111	4	US-09-775-932-18 Sequence 18, Appl
43	85	31.2	121	4	US-09-775-932-14 Sequence 14, Appl
44	85	31.2	128	4	US-09-775-932-12 Sequence 12, Appl
45	85	31.2	149	2	US-08-461-030C-2 Sequence 2, Appl

## ALIGNMENTS

```

RESULT 1
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; EARLIER FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-4

Query Match      46.7%; Score 127; DB 3; Length 142;
Best Local Similarity 50.0%; Pred. No. 4.2e-11;
Matches 24; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY      2 QYKESDCKYHFRFVLKQVOTDHLLEYHLVNMOWTTCOKPETT 49
Db      55 EYKESDCKYVFLVVKTLQALQVNTLLEYLIDVEIRSDCKRPLSTN 102

RESULT 2
US-09-617-302-4
; Sequence 4, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; EARLIER FILING DATE: 2000-07-17
; EARLIER APPLICATION NUMBER: 09/431,480
; EARLIER FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22

```

```

: SOFTWARE: FastSeq for Windows Version 3.0.0
: SEQ ID NO 4
: LENGTH: 142
: TYPE: prt
: ORGANISM: Homo sapiens
US-09-617-302-4

```

Query Match	46.7%;	Score 127;	DB 3;	Length 142;
Best Local Similarity	50.0%;	Pred. No. 4.2e-11;		
Matches 24;	Conservative 12;	Mismatches 12;	Indels 0;	Gaps 0;

```
QY      2 QYNKESDDKHYRIRFVLKVGQVTDHLEHYLNVMQWTTQCKPEYTN 49
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     55 EYNKESDDKYVFLVKTLLQALQVNTNLLLEYLLDVEIARSDCKRPLSTN 102
```

```

US-09-431-480-2
RESULT 3
US-09-431-480-2
; Sequence 2, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Felchhaus, Andrew
; TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-431-480-2

```

Query Match	44.9%;	Score 122;	DB 3;	Length 141;
Best Local Similarity	51.2%;	Pred. No. 2.3e-10;		
Matches 22;	Conservative	9;	Mismatches 12;	Indels 0;
				Gaps 0;

```
QY      2 QYNKESDDKHFRIRFLVKQRÖVTDHLLEYHNLNEMOMTTCK 44
          :|||:|:||: |||:|||:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     54 EYNKASNDLYNFRVVDILKSQEQTSDSLEYYLEVNIA RTMCK 96
```

```

RESULT 4
US-09-617-302-2
Sequence 2, Application US/09617302
Patent No. 6245529
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72 C1
CURRENT APPLICATION NUMBER: US/09/617,302
CURRENT FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/431,480
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 60/109,217
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/156,382
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
US-09-617-302-2

```

Query Match	44.9%	Score 122;	DB 3;	Length 141;
Best Local Similarity	51.2%	Pred. No. 2,3e-10;		
Matches 22; Conservative	9;	Mismatches 12;	Indels 0;	Gaps 0;

```
QY      2 QYNKESDDKYHFRIFRVLKVQRVTDHLHEYLHNEMQMWTTCOK    44  
       :|||:|||::|||::|||::|||::|||::|||::|||::|||  
Db     54 EYNKASNDLYNFRVVDILKSQEITDSLEYLELVENIARTMCKK    96
```

RESULT 5  
US-09-431-480-3  
Semenza 3 Application ITS/09431480

```

APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72
CURRENT APPLICATION NUMBER: US/09/431,480
CURRENT FILING DATE: 1999-11-01
EARLIER APPLICATION NUMBER: 60/109,217
EARLIER FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: 60/156,382
EARLIER FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 142
TYPE: PRT
ORGANISM: Mus musculus
US-09-431-480-3

```

Query Match	44.1%;	Score 120;	DB 3;	Length 142;
Best Local Similarity	43.2%;	Pred. No. 4.5e-10;		
Matches 19;	Conservative 14;	Mismatches 11;	Indels 0;	Gaps 0;

```
QY      2 QYNKESDDKYHFRIFRVLVKQROVTDHLEYHLNVEQMWTTCQKP 45
      :||||:| | : :| : | :| :| :| :| :| :| :| :| :| :|
Db      55 EYNKESEDKYVFLVDKILHAKLQITDRMEYQIDVQISRSNCKP 98
```

RESULT 6  
US-09-617-302-3  
; Sequence 3, Application US/09617302

```

? APPLICANT: Holloway, James L.
? APPLICANT: Feldhaus, Andrew
? TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
?
? FILE REFERENCE: 98-72 C1
? CURRENT APPLICATION NUMBER: US/09/617,302
? CURRENT FILING DATE: 2000-07-17
? PRIOR APPLICATION NUMBER: 09/431,480
? PRIOR FILING DATE: 1999-11-01
? PRIOR APPLICATION NUMBER: 60/109,217
? PRIOR FILING DATE: 1998-11-20
? PRIOR APPLICATION NUMBER: 60/156,382
? PRIOR FILING DATE: 1999-09-28
?
? NUMBER OF SEQ ID NOS: 22
? SOFTWARE: FastSeq for Windows Version 3.0
?
? SEQ ID NO 3
? LENGTH: 142
?
? TYPE: PRT
?
? ORGANISM: Mus musculus
?
? US-09-617-302-3

```

Query Match	44.1%;	Score 120;	DB 3;	Length 142;
Best Local Similarity	43.2%;	Pred. No. 4.5e-10;		
Matches 19;	Conservative 14;	Mismatches 11;	Indels 0;	Gaps 0;

```

QY      2 QYNKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCPK 45
      : : : : : : : : : : : : : : : : : : : : : : : :
DB      55 EYNKESEDKYVFLVDKILHAKLQITDRMEYQIDVQISRSNCKP 98

```

## RESULT 7

US-08-849-303-19

Sequence 19, Application US/08849303

Patent No. 6680424

GENERAL INFORMATION:

APPLICANT: Atkinson, Howard J.

APPLICANT: McPherson, Michael J.

APPLICANT: Urwin, Peter E.

TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS

NUMBER OF SEQUENCES: 79

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Klauber &amp; Jackson

STREET: 411 Hackensack Avenue, 4th Floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/849,303

FILING DATE: 21-May-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Bq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1321-1-003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 127 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEICAL: NO

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

PRIOR FILING DATE: 2000-08-01  
 PRIOR APPLICATION NUMBER: DE 10030149.5  
 PRIOR FILING DATE: 2000-06-20  
 NUMBER OF SEQ ID NOS: 84  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 46  
 LENGTH: 140  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-886-319A-46

US-09-886-319A-46

Query Match 39.7%; Score 108; DB 4; Length 140;  
 Best Local Similarity 43.5%; Pred. No. 2.6e-08;  
 Matches 20; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

```

/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FASTSQ for Windows Version 3.0
/ SEQ ID NO: 5
/ LENGTH: 140
/ TYPE: PR1
/ ORGANISM: Mus musculus
/ OS-09-431-480-5

```

Query Match	38.2%	Score 104	DB 3	Length 140
Best Local Similarity	41.3%	Pred. No. 1e-07		
Matches 19; Conservative	11;	Mismatches 16;	Indels 0;	Gaps 0;

```

QY      2 QYNKESDDKXHFRIFRVLKQGRVTDHLEHLNVEMQWTTCCPEET 47
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      53 EYNGSNDAYHSRAIQVVRARKQLVAGVNYFFDVEMGRITTCRSQT 98

```

```

RESULT 11
US-09-617-302-5
; Sequence 5' Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1995-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-617-302-5

```

Query Match	38.2%;	Score 104;	DB 3;	Length 140;
Best Local Similarity	41.3%;	Pred. No. 1e-07;		
Matches	19;	Conservative	11;	Mismatches 16;
				Indels 0;
				Gaps 0;

```
Oy      2 QYNKESDDKYHFRIFRVLVKQROVTDHLFYLNVEMQMWTQQPBT 47
       :|||:|||:::|||||:|
Db     53 EYNGSNDAYHSRAIQVRARKQLVAGVNYFFDVEMGRITTCFSQT 98
```

12  
 US-08-649-303-18  
 ? Sequence 18, Application US/08849303  
 ? Patent No. 6680424  
 ? GENERAL INFORMATION:  
 ? APPLICANT: Atkinson, Howard J.  
 ? APPLICANT: McPherson, Michael J.  
 ? APPLICANT: Urwin, Peter E.  
 ? TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS  
 ? NUMBER OF SEQUENCES: 79  
 ? CORRESPONDENCE ADDRESS:  
 ? ADDRESS: Klauber & Jackson  
 ? STREET: 411 Hackensack Avenue, 4th Floor  
 ? CITY: Hackensack  
 ? STATE: New Jersey  
 ? COUNTRY: USA  
 ? ZIP: 07601  
 ? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: IBM PC compatible  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: PatentIn Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,303
FILING DATE: 21-MAY-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742.
REFERENCE/DOCKET NUMBER: 1321-1-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ. ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-849-303-18

```

Query Match	38.2%	Score 104	DB 4	Length 140
Best Local Similarity	41.3%	Pred. NO. 1e-07		
Matches 19	Conservative 11	Mismatches 16	Indels 0	Gaps 0

QY            2 QYNKESDDKYHRIFRVLKVQRÖVLDLLEHLNVEMÖWTTCCRPET 47  
               :|||: ||| :||: :||| :||| :||  
DB            53 EYNGSNDAYHSRAIQVVRARKQLVAGVNYPFDVEMGRITTCRSQT 98

RESULT 13  
 US-08-849-303-16  
 Sequence 16, Application US/08849303  
 Patent No. 6680424  
 GENERAL INFORMATION:  
 APPLICANT: Ackinson, Howard J.  
 APPLICANT: McPherson, Michael J.  
 APPLICANT: Uwm, Peter B.  
 TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS  
 NUMBER OF SEQUENCES: 79  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Klauber & Jackson  
 STREET: 411 Hackensack Avenue, 4th Floor  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/849,303  
 FILING DATE: 21-MAY-1997  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jackson Bqd., David A.  
 REGISTRATION NUMBER: 26,742  
 REFERENCE/DOCKET NUMBER: 1321-1-003  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201-487-5800  
 TELEFAX: 201-343-1664  
 TELEX: 133521  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 112 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: NO

Search completed: March 18, 2004, 14:25:37  
Job time : 12.5417 secs

US-08-849-303-16

Query Match	37.9%	Score 103;	DB 4;	Length 112;
Best Local Similarity	40.0%;	Pred. No. 1.1e-07;		
Matches 18;	Conservative 12;	Mismatches 15;	Indels 0;	Gaps 0;

```
QY      2 QYNKESDDKYHFRIFRVLKQROVTDHLLEYHLNVEMQMTCCKPE 46
          ::|||:||:::||:|||||:
Db     26 EFNKRSDAYQSRVVRVVRARKQVSGMNYFLDVELGRITCTKSQ 70
```

RESULT 14  
US-09-775-932-16

; Sequence 16, Application US/097759332

GENERAL INFORMATION:

APPLICANT: University of British Columbia  
TITLE OF INVENTION: Production and use of

FILE REFERENCE: 58069

CURRENT APPLICATION NUMBER: US/09/775,932

CURRENT FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: CA99/00717

PRIOR FILING DATE: 1999-08-05

PRIOR APPLICATION NUMBER: 60/095,503  
PRIOR FILING DATE: 1998-08-05

NUMBER OF SEO ID NOS: 32

```

; NUMBER OF DRY 1D MOD: 32
; SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 16

; LENGTH: 116

TYPE: PRT

; ORGANISM: Gallus sp.  
MS-09-775-932-16

03-03-113-234-16

Query Match	37.9%	Score 103;	DB 4;	Length 116;
Best Local Similarity	44.2%	Pred. No. 1.1e-07;		
Matches 19;	Conservative 10;	Mismatches 14;	Indels 0;	Gaps 0;

QY            2 QYNKESDDKYHFRIFRYLVKQRÖVTDHLLEYHLNVEMÖWTTÇOK 44  
               ::||::|||::|||::|||::|||::|||  
DB            31 EYNRASNDKYSSRVVRVISAKRQLVSGIKYLÖVEIGRTCPK 73

**RESULT 15**

US-09-775-932-24

Sequence 24, Application US/09775932  
Patent No. 6534477

! PALEIL NO. 8534411  
! GENERAL INFORMATION:

APPLICANT: University of British Columbia

**TITLE OF INVENTION:** Production and use of Modified Cystatins

FILE REFERENCE: 58069

CURRENT APPLICATION NUMBER: US/0

PRIOR APPLICATION NUMBER: CA99/00717

PRIOR FILING DATE: 1999-08-05

PRIOR APPLICATION NUMBER: 60/

PRIOR FILING DATE: 1998-08-05

NUMBER OF SE

```

; SOFTWARE: P
; CEO ID NO 34

```

; SEQ ID NO 24  
; LENGTH: 118

TYPE: PRT

**ORGANISM:** *Bos taurus*

US-09-775-932-24

Query Match 37.9%; Score 103; DB 4; Length 118;  
Best Local Similarity 40.0%; Pred. No. 1.2e-07;  
Matches 18; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

```
OY      2 QYNKESDDKYHFRIFRYALQVRÖVTDHLLEYHLNVEMQMTCCKPE    46
       ::|||:::||::|||::|||::|||::|||::|||::|||::|||::
Db     32 EFNKRSDAYQSRRVVRRARQVSGMNYFLDVELGRTCTKSQ        76
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using SW model

Run on: March 18, 2004, 14:15:06 ; Search time 29.6528 Seconds

(without alignments)  
427.913 Million cell updates/sec

Title: US-09-941-314-13

Perfect score: 272  
Sequence: 1 DQYNKESDDKYNHRIFRVLK.....EYHLNVMQWTTCKPETHN 49

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895539 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications AA:\*

- 1: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/prodata/2/pubppa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/prodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/prodata/2/pubppa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/prodata/2/pubppa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/prodata/2/pubppa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/prodata/2/pubppa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/prodata/2/pubppa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/prodata/2/pubppa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/prodata/2/pubppa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/prodata/2/pubppa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/prodata/2/pubppa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/prodata/2/pubppa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/prodata/2/pubppa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/prodata/2/pubppa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/prodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	272	100.0	49	9	US-09-941-314-13
2	272	100.0	115	9	US-09-941-314-3
3	272	100.0	117	9	US-09-941-314-4
4	272	100.0	137	9	US-09-941-314-2
5	266	97.8	80	9	US-09-941-314-15
6	254	93.4	46	9	US-09-941-314-10
7	248	91.2	52	9	US-09-941-314-14
8	154	56.6	59	9	US-09-941-314-16
9	149	54.8	33	9	US-09-941-314-12
10	143	52.6	27	9	US-09-941-314-9
11	143	52.6	35	9	US-09-941-314-8
12	138	50.7	24	9	US-09-941-314-11
13	128	47.1	50	9	US-09-864-761-34822
14	128	47.1	50	9	US-09-864-761-48936
15	114	41.9	145	9	US-09-740-638-2

16	114	41.9	145	13	US-10-006-467-2	Sequence 2, Appl
17	114	41.9	145	14	US-10-235-148-2	Sequence 2, Appl
18	112	41.2	127	8	US-08-849-303-19	Sequence 19, Appl
19	108	39.7	140	14	US-10-376-564-46	Sequence 46, Appl
20	108	39.7	140	14	US-10-376-564-48	Sequence 48, Appl
21	108	39.7	145	14	US-10-168-425-14	Sequence 14, Appl
22	107	39.3	181	15	US-10-264-045-2608	Sequence 2608, Ap
23	104	38.2	140	8	US-08-849-303-18	Sequence 18, Appl
24	103	37.9	112	8	US-08-849-303-16	Sequence 16, Appl
25	103	37.9	116	9	US-09-775-932-16	Sequence 16, Appl
26	103	37.9	118	9	US-09-775-932-24	Sequence 24, Appl
27	103	37.9	120	9	US-09-775-932-2	Sequence 2, Appl
28	103	37.9	139	8	US-08-849-303-15	Sequence 15, Appl
29	103	37.9	136	8	US-09-969-834-4	Sequence 4, Appl
30	103	37.9	146	8	US-08-849-303-17	Sequence 17, Appl
31	103	37.9	146	9	US-09-940-497-3	Sequence 3, Appl
32	103	37.9	146	9	US-09-969-834-3	Sequence 3, Appl
33	103	37.9	146	14	US-10-329-428-3	Sequence 3, Appl
34	103	37.9	146	14	US-10-376-564-47	Sequence 47, Appl
35	94	34.6	165	9	US-09-740-638-5	Sequence 5, Appl
36	94	34.6	165	13	US-10-006-467-5	Sequence 5, Appl
37	94	34.6	165	14	US-10-235-148-5	Sequence 5, Appl
38	93	34.2	48	9	US-09-941-314-17	Sequence 17, Appl
39	93	34.2	121	9	US-09-775-932-8	Sequence 8, Appl
40	93	34.2	141	8	US-08-849-303-24	Sequence 24, Appl
41	93	34.2	141	9	US-09-940-497-6	Sequence 6, Appl
42	90	33.1	111	8	US-08-849-303-26	Sequence 26, Appl
43	89	32.7	111	9	US-09-775-932-18	Sequence 18, Appl
44	85	31.2	121	9	US-09-775-932-14	Sequence 14, Appl
45	85	31.2	128	9	US-09-775-932-12	Sequence 12, Appl

## ALIGNMENTS

RESULT 1  
US-09-941-314-13  
Sequence 13, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-B and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-13

Query Match 100.0%; Score 272; DB 9; Length 49;  
Best Local Similarity 100.0%; Pred. No. 7.2e-29;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKYNHRIFRVLK...EYHLNVMQWTTCKPETHN 49  
DB 1 DQYNKESDDKYNHRIFRVLK...EYHLNVMQWTTCKPETHN 49

RESULT 2  
US-09-941-314-3  
Sequence 3, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-B and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein

FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 115  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-3

Query Match 100.0%; Score 272; DB 9; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1.9e-28;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 49  
DB 30 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 78

## RESULT 3

US-09-941-314-4  
Sequence 4, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 117  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-4

Query Match 100.0%; Score 272; DB 9; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1.9e-28;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 49  
DB 32 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 80

## RESULT 4

US-09-941-314-2  
Sequence 2, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 137  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-2

Query Match 100.0%; Score 272; DB 9; Length 137;  
Best Local Similarity 100.0%; Pred. No. 2.3e-28;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 5

US-09-941-314-15  
Sequence 15, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-15

Query Match 97.8%; Score 266; DB 9; Length 80;  
Best Local Similarity 100.0%; Pred. No. 7.9e-28;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 49  
DB 1 QYNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 48

## RESULT 6

US-09-941-314-10  
Sequence 10, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: ZymoGenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to  
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941.314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 46  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-10

Query Match 93.4%; Score 254; DB 9; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.7e-26;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 49  
DB 1 NKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 46

## RESULT 7

US-09-941-314-14  
Sequence 14, Application US/09941314  
Patent No. US20020142396A1

```
/ GENERAL INFORMATION:
/ APPLICANT: ZymoGenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941,314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14
/ LENGTH: 52
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-14

Query Match          91.2%; Score 248; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 1,2e-25;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 KESDDKYHFRIFRVLKVQROVTDHLEVHLNVEMQWTTCKRPETTN 49
Db      1 KESDDKYHFRIFRVLKVQROVTDHLEVHLNVEMQWTTCKRPETTN 45

RESULT 8
US-09-941-314-16
/ Sequence 16, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: ZymoGenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941,314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 16
/ LENGTH: 59
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-16

Query Match          56.6%; Score 154; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 4,7e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      23 RQVTDHLEVHLNVEMQWTTCKRPETTN 49
Db      1 RQVTDHLEVHLNVEMQWTTCKRPETTN 27

RESULT 9
US-09-941-314-12
/ Sequence 12, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: ZymoGenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941,314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12
/ LENGTH: 33
```

```
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-12

Query Match          54.8%; Score 149; DB 9; Length 33;
Best Local Similarity 100.0%; Pred. No. 1,1e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      24 QVTDHLEVHLNVEMQWTTCKRPETTN 49
Db      1 QVTDHLEVHLNVEMQWTTCKRPETTN 26

RESULT 10
US-09-941-314-9
/ Sequence 9, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: ZymoGenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941,314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 9
/ LENGTH: 27
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-9

Query Match          52.6%; Score 143; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 5,7e-12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DQYNKESDDKYHFRIFRVLKVQROVTD 27
Db      1 DQYNKESDDKYHFRIFRVLKVQROVTD 27

RESULT 11
US-09-941-314-8
/ Sequence 8, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: ZymoGenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941,314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 8
/ LENGTH: 35
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-8

Query Match          52.6%; Score 143; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 7,6e-12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DQYNKESDDKYHFRIFRVLKVQROVTD 27
Db      9 DQYNKESDDKYHFRIFRVLKVQROVTD 35
```

RESULT 12  
US-09-941-314-11  
Sequence 11, Application US/09941314  
Patent No. US20020142396A1  
GENERAL INFORMATION:  
APPLICANT: Zymogenetics, Inc.  
TITLE OF INVENTION: Mammalian Cystatin-B and Its Use to  
FILE REFERENCE: 00-81PC  
CURRENT APPLICATION NUMBER: US/09/941,314  
CURRENT FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: 60/230,230  
PRIOR FILING DATE: 2001-09-01  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-941-314-11

Query Match 50.7%; Score 138; DB 9; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.3e-11;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 ROVTDLHLYHNLVEMQWTTCKRPE 46  
Db 1 ROVTDLHLYHNLVEMQWTTCKRPE 24

RESULT 13  
US-09-864-761-34822  
Sequence 34822, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecm1ca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 34822  
LENGTH: 50  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL109954.10  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2  
OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23  
OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01  
US-09-864-761-34822

Query Match 47.1%; Score 128; DB 9; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DOYNKESDDKXPRIFRVLKVRQ 24  
Db 27 DOYNKESDDKXPRIFRVLKVRQ 50

RESULT 14  
US-09-864-761-48936  
Sequence 48936, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecm1ca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 48936  
LENGTH: 50  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL036677.18  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96  
OTHER INFORMATION: EST HUMAN HIT: A130857.1, EVALUE 5.00e-23  
OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUE 1.00e-01  
US-09-864-761-48936

Query Match 47.1%; Score 128; DB 9; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYKESDDKTHFRFRVLRVQRO 24  
DB 27 DQYKESDDKTHFRFRVLRVQRO 50

RESULT 15  
US-09-740-638-2  
Sequence 2, Application US/09740638  
Patent No. US20020006656A1  
GENERAL INFORMATION:  
APPLICANT: Holloway, James L.  
TITLE OF INVENTION: Zcyas: A Member of the Cystatin  
FILE REFERENCE: 99-104  
CURRENT APPLICATION NUMBER: US/09/740,638  
CURRENT FILING DATE: 2000-12-18  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 145  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-740-638-2

Query Match 41.9%; Score 114; DB 9; Length 145;  
Best Local Similarity 40.4%; Pred. No. 2.8e-07;  
Matches 19; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

QY 3 YNKSDDKTHFRFRVLRVQROVTDHLEHIANEMOWTTCOKPETTN 49  
DB 52 YNKSDDKTHFRFRVLRVQROVTDHLEHIANEMOWTTCOKPETTN 98

Search completed: March 18, 2004, 14:30:54  
Job time : 29.6528 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 9.43056 Seconds  
(without alignments)  
499.799 Million cell updates/sec

Title: US-09-941-314-13  
Perfect score: 272  
Sequence: 1 DQYKESDDKYHFRIFRLVK.....EYHLNVMQWTCQKPEETN 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	44.1	139	2 A45361	cystatin-related e
2	112	41.2	120	2 S10587	cystatin C - rat
3	112	41.2	127	2 S07085	cystatin C precurs
4	104	38.2	140	2 A36163	cystatin C precurs
5	103	37.9	112	1 UD80	cystatin - bovine
6	103	37.9	139	1 UDCH	cystatin precursor
7	103	37.9	146	1 UDHU	cystatin C precursor
8	93	34.2	141	2 B29632	cystatin SA precurs
9	90	33.1	111	2 A28793	cystatin - puff ad
10	82	30.1	141	1 UDHUP2	cystatin SN precurs
11	81	29.8	162	2 A43428	onchocystatin - ne
12	80	29.4	141	1 UDHUP1	cystatin S precurs
13	77	28.3	498	2 T31871	hypothetical prote
14	75	27.6	111	1 JC2040	cystatin - chum sa
15	75	27.6	132	1 JC4918	cystatin precursor
16	75	27.6	133	2 JC4536	cystatin precursor
17	74.5	27.4	142	2 A47142	cystatin D precurs
18	67	24.6	141	2 JQ1470	cystatin S precurs
19	64	23.5	436	1 KGBOL1	kininogen, LMW I P
20	64	23.5	621	1 KGBOL1	kininogen, LMW I P
21	63	23.2	438	2 A47702	glucan 1,3-beta-gl
22	63	23.2	438	2 T52149	beta-glucanase [lm
23	62.5	23.0	448	2 JN0118	glucan 1,3-beta-gl
24	62	22.8	434	1 KGBOL2	kininogen, LMW II
25	62	22.8	619	1 KGBOL2	kininogen, LMW II
26	60	22.1	257	2 T03724	C-type cyclin - xi
27	59	21.7	139	2 B64005	hypothetical prote
28	59	21.7	325	2 F69784	conserved hypochet
29	59	21.7	427	1 KGHUL1	kininogen, LMW pre

30	59	21.7	644	1 KGHUL1	kininogen, LMW pre
31	58.5	21.5	246	2 H90539	ABC transporter at
32	58.5	21.5	382	2 G90540	hypothetical prote
33	58	21.3	47	2 S62326	cystatin C, paroti
34	57.5	21.1	428	2 A45568	highly immunoreact
35	57.5	21.1	1585	2 T19121	probable protein-t
36	57	21.0	288	2 T04401	endonuclease (EC 3
37	56.5	20.8	483	2 S61975	glutathione-disulf
38	56	20.6	200	2 I46051	secreted phosphop
39	56	20.6	588	2 C95252	L-fucose isomerase
40	56	20.6	588	2 A97117	L-fucose isomerase
41	56	20.6	617	2 S19254	nitrate reductase
42	55.5	20.4	164	2 T31026	hypothetical prote
43	55.5	20.4	527	2 E90740	probable enzyme [i
44	55.5	20.4	527	2 G85590	probable enzyme yb
45	55.5	20.4	527	2 G64818	probable membrane

## ALIGNMENTS

### RESULT 1

A45361  
cystatin-related epididymal specific protein - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C/Accession: A45361  
R/Cornwall, G.A.; Orgebin-Chriec, M.C.; Ham, S.R.  
Mol. Endocrinol. 6, 1653-1664, 1992  
A/Title: The CRIS gene: a unique testis-regulated gene related to the cystatin family is  
A/Reference number: A45361; MUID:93078799; PMID:1280328  
A/Accession: A45361  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: nucleic acid  
A/Residues: 1-139 <COR>  
A/Cross-references: GB:549926; NID:g260492; PIDN:AAC35390.1; PID:g260493  
A/Note: sequence extracted from NCBI backbone (NCBIP:118813)  
C/Superfamily: cystatin; cystatin homology  
F:28-139/Domain: cystatin homology <CYS>

Query Match 44.1%; Score 120; DB 2; Length 139;  
Best Local Similarity 43.2%; Pred. No. 3.4e-08;  
Matches 19; Conservative 14; Mismatches 11; Indels 0; Gaps 0;

QY 2 QYKESDDKYHFRIFRLVKQROVTDHLEHLNVMQWTCQKRP 45  
Db 52 EYKESDDKYFLVDKILHAKQITDREYQIDVQSRNCKRP 95

### RESULT 2

S10587  
cystatin C - rat  
C/Species: Rattus sp. (rat)  
C/Date: 21-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 16-Jul-1999  
C/Accession: S10587  
R/Barnard, F.; Barnard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Billard, M.; Gauthu  
Biochem. Biophys. Res. Commun. 161, 161-166, 1990  
A/Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylation  
A/Reference number: S10587; MUID:90380276; PMID:2400577  
A/Accession: S10587  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-120 <ESN>  
A/Note: 43-Asn was also found  
A/Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala  
C/Superfamily: cystatin; cystatin homology  
F:9-120/Domain: cystatin homology <CYS>

Query Match 41.2%; Score 112; DB 2; Length 120;  
Best Local Similarity 43.5%; Pred. No. 3.1e-07;  
Matches 20; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 2 QYKESDDKYHFRIFRLVKQROVTDHLEHLNVMQWTCQKRPET 47



A/Reference number: S48159; MUID:95010016; PMID:7925354  
 A/Accession: S48159  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 24-139 <NDB>  
 R/Lieber, B.; Krieglstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.  
 FEBS Lett. 248, 162-168, 1999  
 A/Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.  
 A/Reference number: S04008; MUID:89252033; PMID:2721673  
 A/Accession: S04008  
 A/Molecule type: protein  
 A/Residues: 97-114 <LAB>  
 R/Colletta, R.; Bird, J.W.C.  
 Gene 130, 175-181, 1993  
 A/Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping  
 A/Reference number: JN0789; MUID:9336172; PMID:8359684  
 A/Accession: JN0789  
 A/Molecule type: DNA  
 A/Residues: 1-139 <CO2>  
 A/Cross-references: GB:M95725  
 A/Note: authors failed to translate the codon for residue 115-Tyr  
 C/Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases  
 C/Genetic:  
 A/Genes: Can  
 A/Introns: 76/3; 114/3  
 C/Superfamily: cystatin: cystatin homology  
 C/Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein  
 F/1-23/Domain: signal sequence #status predicted <SIG>  
 F/24-139/Product: cystatin, long form #status experimental <CYLP>  
 F/30-139/Domain: cystatin homology <CYS>  
 F/32-139/Product: cystatin, short form #status experimental <CYSP>  
 F/76-80/Region: inhibitory #status predicted  
 F/94-104,118-138/Diulfide bonds: #status experimental  
 F/103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 37.9%; Score 103; DB 1; Length 139;  
 Best Local Similarity 44.2%; Pred. No. 5.3e-06;  
 Matches 19; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 2 QYNKESDKYHFRIFVLKVRQVTDHLEHLNEMQWTCOK 44  
 :||:||||:||||:||||:||||:||||:||||:||||:  
 Db 54 EYNRASNDKXSRVRFVIAKRLVSGIKYILQVEIGRTTCK 96

RESULT 7  
 UDHU  
 Cystatin C precursor [validated] - human  
 N/Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma  
 C/Species: Homo sapiens (man)  
 C/Date: 06-Jul-1982 #sequence, revision 31-Mar-1991 #text change 08-Dec-2000  
 C/Accession: S10216; S00004; JI0095; A33400; S02751; A01270; A25434; S12288; A32732; A60  
 R/Abrahamson, M.; Olafsson, I.; Palstedt, A.; Ullsboeck, M.; Lundvall, A.; Jansson, O.  
 Biochem. J. 268, 287-294, 1990  
 A/Title: Structure and expression of the human cystatin C gene.  
 A/Reference number: S10216; MUID:90303202; PMID:2353674  
 A/Accession: S10216  
 A/Molecule type: DNA  
 A/Residues: 1-146 <AB1>  
 A/Cross-references: EMBL:X52255; NID:930257; PIDN:CAA56497.1; PID:g296643  
 R/Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundvall, A.  
 FEBS Lett. 216, 229-233, 1987  
 A/Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the  
 A/Reference number: S00004; MUID:87219149; PMID:3495457  
 A/Accession: S00004  
 A/Molecule type: mRNA  
 A/Residues: 1-146 <AB2>  
 R/Lever, E.; Lopez-Otin, C.; Ghiso, J.; Gellner, D.; Frangione, B.  
 J. Exp. Med. 169, 1771-1778, 1989  
 A/Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a  
 A/Reference number: JI0095; MUID:89235594; PMID:2541223  
 A/Accession: JI0095  
 A/Molecule type: DNA

A/Residues: 1-146 <LEV>  
 A/Cross-references: GB:X61681; NID:930367; PIDN:CAA43856.2; PID:g4490944  
 A/Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredita  
 e)  
 R/Saitoh, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada, H.  
 Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989  
 A/Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which  
 A/Reference number: A33400; MUID:89350949; PMID:2764935  
 A/Accession: A33400  
 A/Molecule type: DNA  
 A/Residues: 1-24, 'T', 26-146 <SA1>  
 A/Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:9181385; PIDN:AA52164.1; PID:g  
 R/Ghiso, J.; Cowan, N.; Frangione, B.  
 Biol. Chem. Hoppe-Seyler 369, 205-208, 1988  
 A/Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron  
 A/Reference number: S02751; MUID:89076507; PMID:3264504  
 A/Accession: S02751  
 A/Molecule type: DNA  
 A/Residues: 82-119 <GH2>  
 A/Cross-references: EMBL:M27769  
 A/Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown i  
 R/Grubb, A.; Lofberg, H.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982  
 A/Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in the  
 A/Reference number: A01270; MUID:82222268; PMID:6283552  
 A/Accession: A01270  
 A/Molecule type: protein  
 A/Residues: 27-111, 'S', 133-146 <GRU>  
 R/Ghiso, J.; Jansson, O.; Frangione, B.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986  
 A/Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland ty  
 A/Reference number: A25434; MUID:86206076; PMID:3517880  
 A/Accession: A25434  
 A/Molecule type: protein  
 A/Residues: 37-93, 'Q', 95-146 <GH1>  
 R/Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Brodtkin, M.; Borchardt, U.; Machleidt, W.  
 Hoppe-Seyler Z. Physiol. Chem. 364, 1487-1496, 1983  
 A/Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatit  
 A/Reference number: S01461; MUID:84110053; PMID:6662458  
 A/Accession: S12288  
 A/Molecule type: protein  
 A/Residues: 27-73 <TUR>  
 R/Brzin, J.; Popovic, T.; Turk, V.  
 Biochem. Biophys. Res. Commun. 118, 103-109, 1984  
 A/Title: Human cystatin, a new protein inhibitor of cysteine proteinases.  
 A/Reference number: A32732; MUID:84128015; PMID:6365094  
 A/Accession: A32732  
 A/Molecule type: protein  
 A/Residues: 27-76 <BR2>  
 R/Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.  
 Scand. J. Clin. Lab. Invest. 50, 85-93, 1990  
 A/Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyste  
 A/Reference number: A60552; MUID:90193615; PMID:2315647  
 A/Accession: A60552  
 A/Molecule type: protein  
 A/Residues: 27-49, 'X', 52-64 <OLA>  
 A/Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d  
 e defective gene is not present in CSF but is found instead in amyloid deposits  
 R/Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.  
 Biol. Chem. Hoppe-Seyler 371, 575-580, 1990  
 A/Title: Different forms of human cystatin C.  
 A/Reference number: S10607; MUID:91025625; PMID:2222856  
 A/Accession: S10607  
 A/Molecule type: protein  
 A/Residues: 27-53 <POP>  
 A/Experimental source: urine, kidney disease  
 A/Note: truncated forms with amino ends at positions 35 and 36 of the precursor were also  
 R/Grubb, A.; Lofberg, H.; Barrett, A.J.  
 FEBS Lett. 170, 370-374, 1984  
 A/Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
 A/Reference number: S01462  
 A/Contents: annotation; disulfide bonds  
 R/Bertl, P.J.; Storer, A.C.

Biochem. J. 302, 411-416, 1994  
 A>Title: Local pH-dependent conformational changes leading to proteolytic susceptibility  
 A/Reference number: S55305; MUID:94379969; PMID:8092991  
 A/Accession: S55305  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 27-49/106-146 <BER>  
 C/Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fluid patients with certain autoimmune diseases.  
 C/Comment: This protein is an inhibitor of cysteine proteinases and may serve an important function.  
 C/Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorrhage.  
 C/Genetics:  
 A/Gene: GDB:CST3  
 A/Cross-references: GDB:119817; OMIM:105150  
 A/Map position: 20p11.2-20p11.2  
 A/Intons: 81/3; 119/3  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline  
 F/1-26/Domain: signal sequence #status predicted <SIG>  
 F/1-27-146/Product: cystatin C #status experimental <MAT>  
 F/35-146/Domain: cystatin homology <CYS>  
 F/81-85/Region: inhibitory #status predicted  
 F/29/Modified site: hydroxyproline (Pro) (partial) #status experimental  
 F/99-109,123-143/Disulfide bonds: #status experimental

Query Match 37.9%; Score 103; DB 1; Length 146;  
 Best Local Similarity 40.0%; Pred. No. 5.6e-06;  
 Matches 18; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDQKXHFRIFFVLKQROVTDHLEHNVEMQWTTCKPE 46  
 Db 59 EYNKASNDMYHSLQVRAKQIVAGVNYFLDVELGRTTCTKQ 103

## RESULT 8

B29632

cystatin SA precursor - human

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1989 #sequence, revision 30-Jun-1989 #text\_change 16-Jul-1999

C/Accession: B29632; S02490; A41422; B27015

R/Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.

Gene 61, 329-338, 1987

A&gt;Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three men

A/Reference number: A91589; MUID:88185836; PMID:3446578

A/Accession: B29632

A/Molecule type: DNA

A/Residues: 1-141 &lt;SAI&gt;

A/Cross-references: GB:M19673; GB:M19170; NID:G186403; PIDN:AAA36116.1; PID:G3186826

A/Note: The authors translated the codon GAC for residue 129 as Asn

R/Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.

Biol. Chem. Hoppe-Seyler 369, 191-197, 1988

A&gt;Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily

A/Reference number: S02489; MUID:89076505; PMID:3202964

A/Accession: S02490

A/Status: not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 21-141 &lt;SA2&gt;

R/Isemura, S.; Saitoh, E.; Sanada, K.

J. Biochem. 102, 693-704, 1987

A&gt;Title: Characterization and amino acid sequence of a new acidic cysteine proteinase in

A/Reference number: A41422; MUID:88139220; PMID:3436950

A/Accession: A41422

A/Molecule type: protein

A/Residues: 25-141 &lt;ISE&gt;

R/Isemura, S.; Saitoh, E.; Sanada, K.; Isemura, M.; Ito, S.

in Cysteine Proteinases and Their Inhibitors, Turk, V., ed., pp.497-505, Walter de Gruyter

A/Title: Cystatin S and the related cysteine proteinase inhibitors in human saliva.

A/Reference number: A27015

A/Accession: B27015

A/Molecule type: protein

A/Residues: 25-134, 'D', 136-141 &lt;IS2&gt;

C/Genetics:

A/Gene: GDB:CST2

A/Cross-references: GDB:119816; OMIM:123856  
 A/Map position: 20p11.2-20p11.2  
 C/Superfamily: cystatin; cystatin homology  
 F/30-141/Domain: cystatin homology <CYS>

Query Match 34.2%; Score 93; DB 2; Length 141;  
 Best Local Similarity 33.3%; Pred. No. 0.00011;  
 Matches 15; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 2 QYNKESDQKXHFRIFFVLKQROVTDHLEHNVEMQWTTCKPE 46  
 Db 54 EYNKATEDYRRLRLVLRARQIVGVNFFPLEVORTICTNSQ 98

## RESULT 9

A28793

cystatin - puff adder

C/Species: Bitis arisanus (puff adder)

C/Date: 15-Dec-1988 #sequence, revision 15-Dec-1988 #text\_change 30-Sep-1993

C/Accession: A28793

R/Rittonja, A.; Evans, H.J.; Machleidt, W.; Barrett, A.J.

Biochem. J. 246, 799-802, 1987

A&gt;Title: Amino acid sequence of a cystatin from venom of the African puff adder (Bitis a

A/Reference number: A28793; MUID:88076861; PMID:3500714

A/Accession: A28793

A/Molecule type: protein

A/Residues: 1-111 &lt;RIT&gt;

C/Superfamily: cystatin; cystatin homology

Query Match 33.1%; Score 90; DB 2; Length 111;  
 Best Local Similarity 38.6%; Pred. No. 0.0002;  
 Matches 17; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 DOYNKESDQKXHFRIFFVLKQROVTDHLEHNVEMQWTTCK 44  
 Db 24 EYNKAGSKNDYFKERVRVAGQGVSVGYIMELTKTTCK 67

## RESULT 10

UDHUP2

cystatin SN precursor [validated] - human

N/Alternate names: cystatin SA-I

C/Species: Homo sapiens (man)

C/Date: 28-May-1986 #sequence, revision 08-Feb-1996 #text\_change 08-Dec-2000

C/Accession: A28110; S02489; A29632; A01273; S19279

R/A1-Haehini, I.; Dickinson, D.P.; Levine, M.J.

J. Biol. Chem. 263, 9381-9387, 1988

A&gt;Title: Purification, molecular cloning, and sequencing of salivary cystatin SA-I.

A/Reference number: A28110; MUID:88243825; PMID:2837486

A/Accession: A28110

A/Molecule type: mRNA

A/Residues: 1-141 &lt;AH&gt;

A/Cross-references: GB:U03870; NID:G337751; PIDN:AAA60299.1; PID:G337752

R/Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.

Biol. Chem. Hoppe-Seyler 369, 191-197, 1988

A&gt;Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily

A/Reference number: S02489; MUID:89076505; PMID:3202964

A/Accession: S02489

A/Status: not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 21-141 &lt;SA2&gt;

R/Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.

Gene 61, 329-338, 1987

A&gt;Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three men

A/Reference number: A91589; MUID:88185836; PMID:3446578

A/Accession: A29632

A/Molecule type: DNA

A/Residues: 1-86, 'I', 88-141 &lt;SAI&gt;

R/Isemura, S.; Saitoh, E.; Sanada, K.

FEBS Lett. 198, 145-149, 1986

A&gt;Title: Characterization of a new cysteine proteinase inhibitor of human saliva, cystati

A/Reference number: A01273; MUID:86164938; PMID:3514272

A/Accession: A01273

A/Molecule type: protein  
 A/Residues: 29-141 <ISE>  
 R/Ramaiahb, N.; Reddy, M.S.; Bergey, E.J.; Harszchky, G.G.; Soni, S.D.; Levine, M.J.  
 Biochem. J. 280, 341-352, 1991  
 A/Title: Large-scale purification and characterization of the major phosphoproteins and  
 A/Reference number: S19279; MUID:92082469; PMID:1747107  
 A/Accession: S19279  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 21-55 <RAM>  
 C/Comment: Human saliva appears to contain several cysteine proteinase inhibitors that a  
 cences. Cystatin SN, with a pI of 7.5, is a much better inhibitor of papain and dipeptidyl  
 C/Genetics:  
 A/Gene: GDB:CST1  
 A/Cross-references: GDB:119815; OMIM:123855  
 A/Map position: 20p11.2-20p11.2  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: cysteine proteinase inhibitor; extracellular protein; saliva  
 F/1-20/Domain: signal sequence #status predicted <SIG>  
 F/21-141/Product: cystatin SA-I #status experimental <MAT>  
 F/28-141/Product: cystatin SN #status experimental <MAT>  
 F/30-141/Domain: cystatin homology <CYS>  
 F/76-80/Region: inhibitory #status predicted  
 F/94-104,118-138/Diulfide bonds: #status predicted

Query Match 30.1%; Score 82; DB 1; Length 141;  
 Best Local Similarity 35.6%; Pred. No. 0.0028;  
 Matches 16; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

OY 2 QYNKESDCKYHFRIFRYLVKQVQVTDHLEHYLVNEMQMTTCQKDE 46  
 DB 54 EYKATDYRRPRLRVLRARQQLVGVNFFPDEVGRTICTKSQ 98

## RESULT 11

A43428  
 onchocystatin - nematode (Onchocerca volvulus)  
 N/Alternate names: cysteine proteinase inhibitor; onchocerciasis antigen  
 C/Species: Onchocerca volvulus  
 C/Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 17-Mar-2000  
 C/Accession: A43428; B43927  
 R/Lustigman, S.; Brozman, B.; Huima, T.; Prince, A.M.; McKerrow, J.H.  
 J. Biol. Chem. 267, 17339-17346, 1992  
 A/Title: Molecular cloning and characterization of onchocystatin, a cysteine proteinase  
 A/Reference number: A43428; MUID:92381053; PMID:1512265  
 A/Accession: A43428  
 A/Status: preliminary  
 A/Molecule type: RNA  
 A/Residues: 1-162 <LUS>  
 A/Cross-references: GB:M37105; NID:g159905; PID:g159906  
 A/Note: sequence extracted from NCBI backbone (NCBIN:111962, NCBI:P.111963)  
 R/Chandraashekar, R.; Maqood, K.; Alvarez, R.M.; Ogunrinade, A.F.; Lujan, R.; Richards Jr  
 J. Clin. Invest. 88, 1460-1466, 1991  
 A/Title: Molecular cloning and characterization of recombinant parasite antigens for imm  
 A/Reference number: A43927; MUID:92042723; PMID:1840605  
 A/Accession: B43927  
 A/Molecule type: mRNA  
 A/Residues: 'P', 37-57, 'R', 73-83, 'N', 85-126, 'W', 128-162 <CHA>  
 A/Cross-references: GB:M60279; NID:g159888  
 A/Experimental source: clone OC 9.3  
 A/Note: sequence extracted from NCBI backbone (NCBIN:65111, NCBI:P.65113)  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: cysteine proteinase inhibitor

Query Match 29.8%; Score 81; DB 2; Length 162;  
 Best Local Similarity 34.1%; Pred. No. 0.0043;  
 Matches 14; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

OY 4 NKESDCKYHFRIFRYLVKQVQVTDHLEHYLVNEMQMTTCQK 44  
 DB 77 NQSDNDEYHMPKILTKVSSQVAGVYKMDVQVARSQCKK 117

## RESULT 12

UDHUP1  
 cystatin S precursor - human  
 N/Alternate names: cystatin SA-III; salivary acidic protein-1  
 C/Species: Homo sapiens (man)  
 C/Date: 25-Feb-1985 #sequence\_revision 08-Feb-1996 #text\_change 16-Jul-1999  
 C/Accession: S17667; S16500; A01272; A29603; S19280; A56608  
 R/Bobek, L.A.; Aguirre, A.; Levine, M.J.  
 Biochem. J. 278, 627-635, 1991  
 A/Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ and  
 A/Reference number: S17667; MUID:91378918; PMID:1898352  
 A/Accession: S17667  
 A/Molecule type: mRNA  
 A/Residues: 1-141 <BOB>  
 A/Cross-references: EMBL:X54667; NID:g30365; PIDN:CA48478.1; PID:g30366  
 R/Lamkin, M.S.; Jensen, U.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.  
 Arch. Biochem. Biophys. 288, 664-670, 1991  
 A/Title: Salivary cystatin SA-III, a potential precursor of the acquired enamel pellicle  
 A/Reference number: S16500; MUID:91378515; PMID:1898055  
 A/Accession: S16500  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 21-134, 'D', 136-141 <IHU>  
 R/Isemura, S.; Saitoh, E.; Sanada, K.  
 J. Biochem. 96, 489-498, 1984  
 A/Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole sal  
 A/Reference number: A91985; MUID:95054716; PMID:6501254  
 A/Accession: A01272  
 A/Molecule type: protein  
 A/Residues: 29-134, 'D', 136-141 <ISB>  
 R/Isemura, S.; Saitoh, E.; Ito, S.; Isemura, M.; Sanada, K.  
 J. Biochem. 96, 1311-1314, 1984  
 A/Title: Cystatin S: a cysteine proteinase inhibitor of human saliva.  
 A/Reference number: A91981; MUID:85104877; PMID:6394600  
 A/Contents: annotation; inhibitor specificity  
 R/Hawke, D.H.; Yuan, P.M.; Wilson, K.J.; Hunkapiller, M.W.  
 Biochem. Biophys. Res. Commun. 145, 1248-1253, 1987  
 A/Title: Identification of a long form of cystatin from human saliva by rapid microbre  
 A/Reference number: A29603; MUID:87270697; PMID:3496880  
 A/Accession: A29603  
 A/Molecule type: protein  
 A/Residues: 21-51 <HAM>  
 R/Ramaiahb, N.; Reddy, M.S.; Bergey, E.J.; Harszchky, G.G.; Soni, S.D.; Levine, M.J.  
 Biochem. J. 280, 341-352, 1991  
 A/Title: Large-scale purification and characterization of the major phosphoproteins and  
 A/Reference number: S19279; MUID:92082469; PMID:1747107  
 A/Accession: S19280  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 21-55 <RAM>  
 R/Johnson, M.; Richardson, C.F.; Bergey, E.J.; Levine, M.J.; Nancollas, G.H.  
 Arch. Oral Biol. 36, 631-636, 1991  
 A/Title: The effects of human salivary cystatins and statherin on hydroxyapatite crystal  
 A/Reference number: A56608; MUID:92074898; PMID:1741693  
 A/Accession: A56608  
 A/Molecule type: protein  
 A/Residues: 21-36 <UDH>  
 A/Note: sequence extracted from NCBI backbone (NCBI:P.67866)  
 C/Comment: authors designate form without phosphate as cystatin S and form containing one ph  
 competitively.  
 C/Genetics:  
 A/Gene: GDB:CST4  
 A/Cross-references: GDB:136381  
 A/Map position: 20p11.2-20p11.2  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: cysteine proteinase inhibitor; phosphoprotein; saliva  
 F/1-20/Domain: signal sequence #status predicted <SIG>  
 F/21-141/Product: cystatin S #status predicted <MAT>  
 F/30-141/Domain: cystatin homology <CYS>  
 F/76-80/Region: inhibitory #status predicted  
 F/94-104,118-138/Diulfide bonds: #status predicted

Query Match 29.4%; Score 80; DB 1; Length 141;  
 Best Local Similarity 33.3%; Pred. No. 0.005;  
 Matches 15; Conservative 12; Mismatches 18; Indels 0; Gaps 0;  
 QY 2 QYNKESDDKXHFRIFRYLKVGROVTDHLEHNLVEMQWTTGCKPE 46  
 DB 54 EYNKATEDEYRRLPLQVRAKARQTFGGVNFDFVEVGRTITKSKQ 98

## RESULT 13

T31871  
 hypothetical protein F41E6.6 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C/Accession: T31871  
 R/Sammons, L.; Murray, J.  
 submitted to the EMBL Data Library, July 1997  
 A/Description: The sequence of C. elegans cosmid F41E6.  
 A/Reference number: Z21095  
 A/Accession: T31871  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-498 <SAM>  
 A/Cross-references: EMBL:AF016448; PIRN:AAB65956.1; GSPDB:GN00023; CESP:F41E6.6  
 A/Experimental source: Strain Bristol N2; clone F41E6  
 C/Genetics:  
 A/Map position: 5  
 A/Intons: 34/3; 102/3; 140/2; 180/2; 336/2; 357/2; 417/1

Query Match 28.3%; Score 77; DB 2; Length 498;  
 Best Local Similarity 27.7%; Pred. No. 0.05;  
 Matches 13; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

QY 1 DOYNKESDDKXHFRIFRYLKVGROVTDHLEHNLVEMQWTTGCKPET 47  
 DB 42 EKNQGSNDAYKWEIDRTWEVERKLSGINSIFVTLYKIDCKKQOT 88

## RESULT 14

JC2040  
 cystatin - chum salmon  
 N/Alternate names: cysteine proteinase inhibitor  
 C/Species: Oncorhynchus keta (chum salmon)  
 C/Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 16-Jul-1999  
 C/Accession: JC2040  
 R/Koide, Y.; Noso, T.  
 Biosci. Biotechnol. Biochem. 58, 164-169, 1994  
 A/Title: The complete amino acid sequence of pituitary cystatin from chum salmon.  
 A/Reference number: JC2040; MUID:94162738; PMID:7764512  
 A/Accession: JC2040  
 A/Molecule type: protein  
 A/Residues: 1-111 <KOI>  
 C/Comment: The intracellular role of this protein is the inhibition of intralysosomal pr  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: cysteine proteinase inhibitor  
 F/2-111/Domain: cystatin homology <CYS>  
 F/48-52/Region: inhibitory  
 F/89-109/Disulfide bonds: #status experimental

Query Match 27.6%; Score 75; DB 1; Length 111;  
 Best Local Similarity 30.2%; Pred. No. 0.017;  
 Matches 13; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRYLKVGROVTDHLEHNLVEMQWTTGCK 44  
 DB 26 EHNKKTNDMFVRQYAKVNAQKQVSGMKYIFTVQMGRTPCRK 68

## RESULT 15

JC4918  
 cystatin precursor - chum salmon  
 C/Species: Oncorhynchus keta (chum salmon)

C/Date: 26-Sep-1996 #sequence\_revision 01-Nov-1996 #text\_change 21-Jan-2000  
 C/Accession: JC4918  
 R/Yamashita, M.; Komasawa, S.  
 J. Biochem. 120, 483-487, 1996  
 A/Title: Molecular cloning and gene expression of chum salmon cystatin.  
 A/Reference number: JC4918; MUID:97058285; PMID:8902609  
 A/Accession: JC4918  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-132 <YAM>  
 A/Cross-references: DDBJ:DB6628  
 A/Experimental source: liver  
 C/Comment: This protein is a homolog of mammalian cystatin C, and acts as an extracellular  
 C/Superfamily: cystatin; cystatin homology  
 C/Keywords: liver  
 F/1-21/Domain: signal sequence #status predicted <SIG>  
 F/23-132/Domain: cystatin homology <CYS>

Query Match 27.6%; Score 75; DB 2; Length 132;  
 Best Local Similarity 30.2%; Pred. No. 0.02;  
 Matches 13; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRYLKVGROVTDHLEHNLVEMQWTTGCK 44  
 DB 47 EHNKKTNDMFVRQYAKVNAQKQVSGMKYIFTVQMGRTPCRK 89

Search completed: March 18, 2004, 14:23:17  
 Job time : 9.43056 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 5.8333 Seconds  
(without alignments)  
437.389 Million cell updates/sec

Title: US-09-941-314-13

Perfect score: 272  
Sequence: 1 DQYNKSDDKCHFRIRVVK.....EYHLNVMQMTQCKPPTTN 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues  
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	272	100.0	CS11_HUMAN	Q9H112 homo sapien
2	177	65.1	CS11_MOUSE	Q9D126 mus musculu
3	127	46.7	CS18_HUMAN	O60676 homo sapien
4	120	44.1	CS18_MOUSE	P32766 mus musculu
5	115	42.3	CS18_RAT	O89969 ratu muscu
6	112	41.2	CS18_RAT	P14841 ratu muscu
7	110	40.4	CS18_RAT	O97862 ratu muscu
8	108	39.7	CS18_MOUSE	P21460 mus musculu
9	105	38.6	CS18_MOUSE	O19093 salm mri sci
10	103	37.9	CS18_MOUSE	P01038 gallu galli
11	103	37.9	CS18_MOUSE	P01034 homo sapien
12	103	37.9	CS18_MOUSE	O19092 macaca mula
13	103	37.9	CS18_MOUSE	P01035 bos tauru
14	99	36.4	CS18_MOUSE	P81061 cotu muscu
15	94	34.6	CS18_MOUSE	O9H116 homo sapien
16	93	34.2	CS18_MOUSE	P09228 homo sapien
17	90	33.1	CS18_MOUSE	P08935 btiis ariet
18	89	32.7	CS18_MOUSE	P35481 cypitrus ca
19	85	31.2	CS18_MOUSE	O15828 homo sapien
20	82	30.1	CS18_MOUSE	P01037 homo sapien
21	81	29.8	CS18_MOUSE	P22085 onchocerca
22	80	29.4	CS18_MOUSE	P01036 homo sapien
23	75	27.6	CS18_MOUSE	O9H119 onchocerca
24	75	27.6	CS18_MOUSE	P28322 homo sapien
25	74.5	27.4	CS18_MOUSE	O76096 homo sapien
26	71	26.1	CS18_MOUSE	P19313 ratu muscu
27	67	24.6	CS18_MOUSE	O89099 mus musculu
28	64	23.5	CS18_MOUSE	P01046 bos tauru
29	64	23.5	CS18_MOUSE	P01044 bos tauru
30	64	23.5	CS18_MOUSE	P29717 candida alb
31	63	23.0	CS18_MOUSE	P23776 saccharomyc
32	62.5	23.0	CS18_MOUSE	P01047 bos tauru
33	62	22.8	CS18_MOUSE	P01047 bos tauru

34	62	22.8	619	1	KNH2_BOVIN	P01045 bos tauru
35	61.5	22.6	137	1	CS18_MOUSE	O92016 mus musculu
36	60	22.1	99	1	CYT_NAIRT	P81714 naia atra (
37	60	22.1	257	1	CGIC_ORISA	P93411 oryza sativ
38	60	22.1	661	1	KNH_MOUSE	O08677 mus musculu
39	59	21.7	644	1	KNH_HUMAN	P01042 homo sapien
40	57.5	21.1	2200	1	LAR_CAEEL	O9bms caenorhabdi
41	56.5	20.8	483	1	GSRR_YEAST	P41921 saccharomyc
42	56	20.6	200	1	SP24_BOVIN	O27967 bos tauru
43	56	20.6	621	1	NIAI_MAIZ	P17571 zea mays (m
44	55.5	20.4	527	1	YBIP_ECOLI	P75785 escherichia
45	55	20.2	525	1	HRG_HUMAN	P04196 homo sapien

## ALIGNMENTS

RESULT 1  
ID CS11\_HUMAN STANDARD; PRT; 137 AA.  
AC O9H112; O9H113;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE Cystatin II precursor.  
GN CS11 OR CS18.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N. A. (ISOFORMS 1 AND 2).  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.B., Connor R.E., Cobby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.B., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Leiva-Salido M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurtry A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Suleston J.B.,  
RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Walli M., Wallis J.M.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RL "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:665-871(2001).  
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=O9H112-1; Sequence=Displayed;  
CC Name=2;  
CC Note=O9H112-2; Sequence=VSP\_001260;  
CC Note=NO experimental confirmation available;  
CC -!- SIMILARITY: Belongs to the cystatin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL: AL096677; CAC13170.1; -  
DR EMBL: AL096677; CAC17423.1; -  
DR HSSP: P01038; 1490.  
DR Genew: HGNC:15959; CST11.  
DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; Cystatin; 1.  
DR SMART: SM00043; CY, 1.  
DR PROSITE: PS00287; CYSTATIN; FALSE\_NEG.  
DR Thiol protease inhibitor; Signal; Alternative splicing.  
KW SIGNAL 1 25  
FT CHAIN 26 137 CYSTATIN 11.  
FT SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).  
FT DISULFID 93 101 BY SIMILARITY.  
FT CARBOHYD 114 134 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPIC 76 110 Missing (in isoform 2).  
FT /FTId=VSP\_001260.  
SQ SEQUENCE 137 AA; 16375 MW; C5856C8C39A585C3B CRC64;

Query Match 100.0%; Score 272; DB 1; Length 137;  
Best Local Similarity 100.0%; Pred. No. 6,3e-28;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQVKKSDDKYHFRIPVLYKQVQVTDHLEHYNEMQWTCCKPEPTN 49  
Db 52 DQVKKSDDKYHFRIPVLYKQVQVTDHLEHYNEMQWTCCKPEPTN 100

## RESULT 2

CS11\_MOUSE STANDARD; PRT; 139 AA.

AC Q9D269;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cystatin, 11 precursor.  
GN CST11.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spididymis;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
RA Aizawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baren G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gusticich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,  
RA Wyszynski A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
CC - SUBCELLULAR LOCATION: Secreted (Potential).  
CC - SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL: AK020300; BAB32061.1; -  
DR HSSP: P01034; 1696.  
DR MED: MGI:1925490; Cst11.  
DR InterPro: IPR000010; Cystatin.  
DR Pfam: PF00031; Cystatin; 1.  
DR SMART: SM00043; CY, 1.  
DR PROSITE: PS00287; CYSTATIN; FALSE\_NEG.  
KW Thiol protease inhibitor; Signal.  
KW SIGNAL 1 28  
FT CHAIN 29 139 CYSTATIN 11.  
FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).  
FT DISULFID 94 102 BY SIMILARITY.  
FT CARBOHYD 115 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT /FTId=VSP\_001260.  
SQ SEQUENCE 139 AA; 16217 MW; F228D9815FA32640 CRC64;

Query Match 65.1%; Score 177; DB 1; Length 139;  
Best Local Similarity 64.6%; Pred. No. 9,4e-16;  
Matches 31; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DQVKKSDDKYHFRIPVLYKQVQVTDHLEHYNEMQWTCCKPEPTN 48  
Db 53 DQVKKSDDKYHFRIPVLYKQVQVTDHLEHYNEMQWTCCKPEPTN 100

## RESULT 3

CST8\_HUMAN STANDARD; PRT; 142 AA.

AC O60676;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin  
DE 8).  
GN CST8 OR CRBS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=95344753; PubMed=7619504;  
RA Cornwell G.A., Hann S.R.,  
RT "Transient appearance of CRBS protein during spermatogenesis and  
RT caput epididymal sperm maturation."  
RL Mol. Reprod. Dev. 41:37-46(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dham P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.R., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehesvahti M.H., Levenha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parfery A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramday H.,  
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showgreen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epithelium.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF059244; AAC14707.1; -;  
 CC EMBL: AL109954; CAB64234.1; -;  
 CC HSSP: P01034; 1G96.  
 DR GeneW: HNCN:2480; CST8.  
 DR GO: GO:0004869; F:Cysteine protease inhibitor activity; TAS.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; Cystatin; 1.  
 DR SMART: SM00043; CY: 1.  
 DR Thiol protease inhibitor; Signal; Polymorphism.  
 KW Thiol protease inhibitor; Signal; Polymorphism.  
 FT SIGNAL 1 21  
 FT CHAIN 22 142  
 FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 95 105 BY SIMILARITY.  
 FT DISULFID 119 139 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 142 142 A -> P (in dbSNP:1054633).  
 FT SIGNAL 142 142 /FTID=VAR 014527.  
 SQ SEQUENCE 142 AA; 16275 MW; 9A3512757E0F4ECD CRC64;  
 Query Match 46.7%; Score 127; DB 1; Length 142;  
 Best Local Similarity 50.0%; Pred. No. 2.5e-09;  
 Matches 24; Conservative 12; Mismatches 12; Indels 0; Gaps 0;  
 QY 2 QYNKESDCKYHFRFLVLYKQROVTDHLEHLNEMQWTTQCKRP 49  
 Db 55 EYNKESDKYVFLVYKTIQAOQVNTNLELYLDIVEIARSDCKRPLSTN 102  
 RESULT 4  
 CST8\_MOUSE STANDARD; PRT; 142 AA.  
 AC P32766; O89102;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cystatin-related epididymal spermato-genic protein precursor (Cystatin-  
 DE related epididymal specific protein) (Cystatin 8).  
 GN CST8 OR CRES.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CH, and CD-1;  
 RX MEDLINE=9247899; PubMed=10229662;  
 RA Cornwall G.A., Hsia N., Sutton H.G.;  
 RT "Structure, alternative splicing and chromosomal localization of the  
 RT cystatin-related epididymal spermato-genic gene";  
 RL Biochem. J. 340:85-93(1999).  
 RN [2]  
 RP SEQUENCE OF 4-142 FROM N.A.  
 RC TISSUE=Epididymis;  
 RX MEDLINE=93078799; PubMed=1280328;  
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;  
 RT "The CRES gene: a unique testis-regulated gene related to the cystatin  
 RT family is highly restricted in its expression to the proximal region  
 RT of the mouse epididymus";  
 RL Mol. Endocrinol. 6:1653-1664(1992).  
 CC -1- FUNCTION: Performs a specialized role during sperm development and  
 CC maturation.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower  
 CC expression in the testis. Within the testis it is localized to the  
 CC elongating spermatids, whereas within the epididymis it is  
 CC exclusively synthesized by the proximal caput epithelium.  
 CC -1- INDUCTION: Testicular factors or hormones other than androgens  
 CC present in the testicular fluid may be involved in the regulation  
 CC of CRES gene expression.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF091503; AAC61754.1; -;  
 CC EMBL: AF090691; AAC36316.1; -;  
 CC EMBL: S4926; AAC35390.1; -;  
 CC PIR: A45361; A45361.  
 DR HSSP: P01034; 1G96.  
 DR MGD: MGI:107161; Cst8.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; Cystatin; 1.  
 DR SMART: SM00043; CY: 1.  
 KW Thiol protease inhibitor; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 142  
 FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).  
 FT DISULFID 95 105 BY SIMILARITY.  
 FT DISULFID 119 139 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 4 15 PMLSILRPIIP -> GTREQVGSBSK (IN REF. 2).  
 SQ SEQUENCE 142 AA; 16288 MW; 50B446B98F6673E CRC64;  
 Query Match 44.1%; Score 120; DB 1; Length 142;  
 Best Local Similarity 43.2%; Pred. No. 1.9e-08;  
 Matches 19; Conservative 14; Mismatches 11; Indels 0; Gaps 0;  
 QY 2 QYNKESDCKYHFRFLVLYKQROVTDHLEHLNEMQWTTQCKRP 45  
 Db 55 EYNKESDKYVFLVYKTIHAKLQITDMEYQIDVOISRSNCKRP 98  
 RESULT 5  
 CST8\_RAT STANDARD; PRT; 142 AA.  
 AC O88969;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

```

DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
DE 8).
GN CSF8 OR CRFS.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Epididymis;
RX MEDLINE=99247899; PubMed=10229682;
RA Cornwell G.A., Hsia N., Sulton H.G.;
RT "Structure, alternative splicing and chromosomal localization of the
RT cystatin-related epididymal spermatogenic gene.";
RL Biochem. J. 340:85-93(1999).
CC -1- FUNCTION: Performs a specialized role during sperm development and
CC maturation.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF090692; AAC6317.1; -.
DR HSBP, P01034; I696.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin, 1.
DR SMART; SM00043; CY; 1.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
FT PROTEIN.
FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
FT DISULFID 95 105 BY SIMILARITY.
FT DISULFID 119 139 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAc. . .) (POTENTIAL).
SQ SEQUENCE 142 AA; 16246 MW; F8B73FPA6B6CB34 CRC64;

Query March 42.3%; Score 115; DB 1; Length 142;
Best Local Similarity 43.2%; Pred. No. 8.5e-08;
Matches 19; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

OY 2 QYNKESDCKYHFRIFVLKVGQVTDHLEHLNEMQMTTCQKP 45
Db 55 EYNGSEBDKYFLDLKTLHATLQITDRMEYHIDVQISRSNCRKP 98

RESULT 6
CYTC_RAT STANDARD; PRT; 127 AA.
ID CYTC_RAT
AC P14841;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor (Fragment).
GN CS73.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Buffalo;
RX MEDLINE=90091212; PubMed=2689174;
RA Cole T., Dickson P.W., Esnard F., Averill F., Risbridger G.,
RA Gauthier F., Schneider G.;
RT "The cDNA structure and expression analysis of the genes for the

```

```

RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
RT in rat brain.";
RL Eur. J. Biochem. 186:35-42(1989).
RN [2]
RP SEQUENCE OF 8-127.
RX MEDLINE=90380276; PubMed=2400577;
RA Esnard F., Esnard A., Faucher D., Capony J.-P., Derancourt J.,
RA Brillard M., Gauthier F.;
RT "Rat cystatin C: the complete amino acid sequence reveals a site for
RT N-glycosylation.";
RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).
RN [3]
RP SEQUENCE OF 8-49.
RX MEDLINE=88313020; PubMed=3044831;
RA Esnard A., Esnard F., Faucher D., Gauthier F.;
RT "Two rat homologues of human cystatin C.";
RL FEBS Lett. 236:475-478(1988).
RN [4]
RP SEQUENCE OF 8-20.
RC TISSUE=Sertoli cells;
RX MEDLINE=9225121; PubMed=1563513;
RA Esnard A., Esnard F., Guillon F., Gauthier F.;
RT "Production of the cysteine proteinase inhibitor cystatin C by rat
RT Sertoli cells.";
RL FEBS Lett. 300:131-135(1992).
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity. Known to inhibit cathepsin B,
CC H. and L.
CC -----
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X16957; CAA34831.1; -.
DR PIR; S07085; S07085.
DR PIR; S10587; S10587.
DR HSBP, P01034; I696.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin, 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Signal.
FT NON_TER 1 1
FT SIGNAL 8 127 CYSTATIN C.
FT CHAIN 18 18 REACTIVE SITE.
FT ACT SITE 62 66 SECONDARY AREA OF CONTACT.
FT SITE 80 90 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 125 25 A -> E (IN REF. 2).
FT CONFLICT 25 25
SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;

Query March 41.2%; Score 112; DB 1; Length 127;
Best Local Similarity 43.5%; Pred. No. 1.8e-07;
Matches 20; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

OY 2 QYNKESDCKYHFRIFVLKVGQVTDHLEHLNEMQMTTCQKPE 47
Db 40 EYNGSNDAYHSRAIQVYRARKQVLVAGINYYLDVEMERTTCTSGT 85

RESULT 7
CYTC_RABIT STANDARD; PRT; 148 AA.
ID CYTC_RABIT
AC O97862;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```



```

SQ SEQUENCE 140 AA; 15531 MW; 3A563406DD58D0F5 CRC64;
Query Match          39.7%; Score 108; DB 1; Length 140;
Best Local Similarity 43.5%; Pred. No. 6.6e-07;
Matches 20; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

OY      2 QYNKESDDKYHPRIFVLKVQROVDLHEYLHNVEMQWTCOKPET 47
       :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|
DB      53 EYNKASNDMYHSRALQVVRARQLAVAGNVPFLDVEMGTTCTKSQT 98

RESULT 9
CYTC_SAISC
ID_CYT_C SAISC STANDARD; PRT; 146 AA.
AC OI9093;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor..
GN CS73.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_Taxid=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA "Wei L.H., Walker L.C., Levy E.;
RT 'Cystatin C. Icelandic-like mutation in an animal model of
RL cerebrovascular beta-amyloidosis.';
RT Stroke 27:2080-2085(1996)
CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -!- SIMILARITY: Belongs to the cystatin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
-----
DR EMBL; U52028; AAB64051.1; -.
DR HSSP; P01034; I396.
DR InterPro; IPRO00010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 1 146 BY SIMILARITY.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 15946 MW; 08196353C0306AA3 CRC64;

Query Match          38.6%; Score 105; DB 1; Length 146;
Best Local Similarity 44.2%; Pred. No. 1.7e-06;
Matches 19; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

OY      2 QYNKESDDKYHPRIFVLKVQROVDLHEYLHNVEMQWTCOK 44
       :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|
DB      59 EYNKASNDMYHSRALQVVRARQLAVAGNVPFLDVEMGTTCTCK 101

RESULT 10
CYT_CHICK
ID_CYT_CHICK STANDARD; PRT; 139 AA.
NC P01038;
DT 21-JUL-1986 (Rel. 01, Created)
```

DT 01-OCT-1989 (Rel. 12, last sequence update)  
DT 28-FEB-2003 (Rel. 41, last annotation update)  
DE Cystatin precursor (Egg-white cystatin).  
OS Gallus gallus (Chicken).  
OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxId=9031.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90008973; PubMed=2793849;  
RA Coellella R., Sakaguchi Y., Nagase H., Bird J.W.C.;  
RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,  
RT and tissue distribution."  
RL J. Biol. Chem. 264:117164-117169(1989).  
RN [2]  
RP SEQUENCE OF 24-139.  
RX MEDLINE=84178305; PubMed=6712597;  
RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;  
RT "Cystatin. Amino acid sequence and possible secondary structure."  
RL Biochem. J. 217:813-817(1984).  
RN [3]  
RP SEQUENCE OF 24-139.  
RX MEDLINE=84110059; PubMed=6662498;  
RA Turk V., Brzin J., Longner M., Riltonja A., Eropkin M., Borchart U.,  
RA Machleidt W.;  
RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
RT of cystatin from chicken egg white."  
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).  
RN [4]  
RP CHARACTERIZATION OF PROTEIN.  
RX MEDLINE=83256421; PubMed=6409085;  
RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,  
RA Sunter D.C., Barrett A.J.;  
RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved  
RT purification from egg white, characterization, and detection in  
RT chicken serum."  
RL Biochem. J. 211:129-138(1983).  
RN [5]  
RP DISULFIDE BONDS.  
RA Grubb A., Loeffberg H., Barrett A.J.;  
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
RT cystatin."  
RL FEBS Lett. 170:370-374(1984).  
RN [6]  
RP PHOSPHORYLATION.  
RX MEDLINE=89252033; PubMed=2721673;  
RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,  
RA Bode W.;  
RT "The cysteine proteinase inhibitor chicken cystatin is a  
RT phosphoprotein."  
RL FEBS Lett. 248:162-168(1989).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=89052676; PubMed=3191914;  
RA Bode W., Engst R., Musil D., Thiele U., Huber R., Karsnikov A.,  
RA Brzin J., Kos J., Turk V.;  
RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and  
RT its possible mode of interaction with cysteine proteinases."  
RL EMBO J. 7:2593-2599(1988).  
RN [8]  
RP STRUCTURE BY NMR.  
RX MEDLINE=94087719; PubMed=8263912;  
RA Dickmann T., Mitschang L., Hofmann M., Kos J., Turk V.,  
RA Auerwald E.A., Jeanicke R., Oeschkat H.;  
RT "The structures of native phosphorylated chicken cystatin and of a  
RT recombinant unphosphorylated variant in solution."  
RL J. Mol. Biol. 234:1049-1059(1993).  
CC -1- FUNCTION: This protein binds tightly to and inhibits a variety of  
CC thiol processes including ficin, papain, and cathepsins B, C, H,  
CC and L. Although isolated from egg white, it is also present in  
CC serum.  
CC -1- SIMILARITY: Belongs to the cystatin family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 CC EMBL J05077; AAA48744.1; -  
 DR PIR; A34456; UDCH.  
 DR PDB; 1A67; 31-JAN-94.  
 DR PDB; 1A67; 27-MAY-98.  
 DR PDB; 1A90; 17-JUN-98.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 KM Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 139  
 FT ACT SITE 32 32 CYSTATIN.  
 FT SITE 76 80 REACTIVE SITE.  
 FT DISULFID 94 104 SECONDARY AREA OF CONTACT.  
 FT DISULFID 118 138  
 FT MOD\_RES 103 103 PHOSPHORYLATION (PARTIAL).  
 FT STRAND 35 36  
 FT TURN 39 40  
 FT TURN 42 51  
 FT TURN 52 52  
 FT HELIX 53 56  
 FT HELIX 57 58  
 FT STRAND 63 77  
 FT STRAND 81 95  
 FT TURN 96 97  
 FT TURN 99 100  
 FT HELIX 101 108  
 FT STRAND 115 125  
 FT TURN 126 129  
 FT STRAND 130 139  
 SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;  
 Query Match 37.94; Score 103; DB 1; Length 139;  
 Best Local Similarity 44.24; Pred. No. 2.9e-06;  
 Matches 19; Conservative 10; Mismatches 14; Indels 0; Gaps 0;  
 QY 2 QYNKESDDKYHFRIFRYLVKQVQVTDHLEVHLNVMQWTCOK 44  
 :||:|||||:|||||:|||||:|||||:|||||  
 54 EYNRASNDKYSRVRVYISAKRLQVSGIKIYLVEIGRTTCPK 96  
 DB  
 RESULT 11  
 CYTC\_HUMAN STANDARD; PRT; 146 AA.  
 AC P01034;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)  
 DE (Post-gamma-globulin).  
 GN CST3  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=87219149; PubMed=3495457;  
 RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;  
 RT "Molecular cloning and sequence analysis of cDNA coding for the  
 RT precursor of the human cysteine proteinase inhibitor cystatin C.";  
 RL FEBS Lett. 216:229-233(1987).  
 -----

-----  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=90303202; PubMed=2363674;  
 RA Abrahamson M., Olafsson I., Palsdottir A., Ulfvack M., Lundwall A.,  
 RA Jansson O., Grubb A.;  
 RT "Structure and expression of the human cystatin C gene";  
 RL Biochem. J. 268:287-294(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A. (HCHWA VARIANT).  
 RC TISSUE=Brain;  
 RX MEDLINE=89235594; PubMed=2541223;  
 RA Levy E., Lopez-Otin C., Ghiso J., Gellner D., Frangione B.;  
 RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is  
 RT related to a mutation in the cystatin C gene, an inhibitor of  
 RT cysteine proteases.";  
 RL J. Exp. Med. 169:1771-1778(1989).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89350949; PubMed=2764935;  
 RA Satoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,  
 RA Iemura S., Sanada K.;  
 RT "The human cystatin C gene (CST3) is a member of the cystatin gene  
 RT family which is localized on chromosome 20.";  
 RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Dickinson D.P., Hewett-Emmett D., Thiesse M.;  
 RT "Acquisition of complex patterns of differential expression in  
 RT epithelial cell populations during the evolution of type 2 cystatin  
 RT genes.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.W., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.B., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehesvahto M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McComachie L.J., McElay K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Rose M.T., Scott C.B., Sehra H.K., Showkhen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wali M., Wallis J.M., Williams S.A.,  
 RA Whitehead S.L., Whiteaker P., Willey D.L., Williams L., Williams S.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marushka K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 -----

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEvan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallue D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [8]  
 RP SEQUENCE OF 27-146.  
 RX MEDLINE=8222268; PubMed=6283552;  
 RA Grubb A., Loeffberg H.,  
 RT "Human gamma-trace, a basic microprotein: amino acid sequence and  
 RT presence in the adenohypophysis.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027(1982).  
 RN [9]  
 RP SEQUENCE OF 27-73.  
 RX MEDLINE=84110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longner M., Riconja A., Eropkin M., Borchart U.,  
 RA Mechleide W.,  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 RT of cystatin from chicken egg white.",  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1993).  
 RN [10]  
 RP SEQUENCE OF 27-76.  
 RX MEDLINE=84128015; PubMed=6365094;  
 RA Brzin J., Popovic T., Turk V.,  
 RT "Human cystatin, a new protein inhibitor of cysteine proteinases.",  
 RL Biochem. Biophys. Res. Commun. 118:103-109(1984).  
 RN [11]  
 RP DISULFIDE BONDS.  
 RA Grubb A., Loeffberg H., Barrett A.J.,  
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
 RT cystatin.",  
 RL FEBS Lett. 170:370-374(1984).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.  
 RX MEDLINE=21173909; PubMed=11276250;  
 RA Janowski R., Kozak M., Jankowska E., Grzonka Z., Grubb A.,  
 RA Abramson M., Jaskolski M.,  
 RT "Human cystatin C, an amyloidogenic protein, dimerizes through  
 RT three-dimensional domain swapping.",  
 RL Nat. Struct. Biol. 8:316-320(2001).  
 RN [13]  
 RP VARIANT GLN-94.  
 RX MEDLINE=92316504; PubMed=1352269;  
 RA Abramson M., Jonsdottir S., Olafsson I., Jensen O., Grubb A.,  
 RT "Hereditary cystatin C amyloid angiopathy: identification of the  
 RT disease-causing mutation and specific diagnosis by polymerase chain  
 RT reaction based analysis.",  
 RL Hum. Genet. 89:377-380(1992).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 CC thought to serve an important physiological role as a local  
 CC regulator of this enzyme activity.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,  
 CC vas deferens, brain, thymus, and ovary and the lowest in the  
 CC submandibular gland.  
 CC -1- DISEASE: Defects in CSN3 are a cause of hereditary cerebral  
 CC hemorrhage with amyloidosis (HCHWA) [MIM:105150], also known as  
 CC cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis  
 CC Icelandic type. HCHWA is characterized by a thickening of the  
 CC cerebral arteriole walls with deposition of material with the  
 CC characteristics of amyloid.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC DR EMBL; X05607; CAA29096.1; -;  
 CC DR EMBL; X52255; CAA36497.1; -;  
 CC DR EMBL; M27891; AAA52164.1; -;  
 CC DR EMBL; M27889; AAA52164.1; JOINED.  
 CC DR EMBL; M27890; AAA52164.1; JOINED.  
 CC DR EMBL; X61681; CAA43856.2; -;  
 CC DR EMBL; X61682; CAA43856.2; JOINED.  
 CC DR EMBL; X61683; CAA43856.2; JOINED.  
 CC DR EMBL; AF319564; AAK11570.1; -;  
 CC DR EMBL; AL121894; CAC05424.1; -;  
 CC DR EMBL; BC013083; AAK13083.1; -;  
 CC DR PIR; S10216; UDHU.  
 CC DR PDB; 1G96; 06-APR-01.  
 CC DR Genew; HGNC:2475; CST3.  
 CC DR MIM; 604312; -;  
 CC DR MIM; 105150; -;  
 CC DR InterPro; IPR000010; Cystatin.  
 CC DR Pfam; PF00043; Cyt. 1.  
 CC DR SMART; SM00043; Cyt. 1.  
 CC DR PROSITE; PS00287; CYSTATIN; 1.  
 CC KM Thiol protease inhibitor; Amyloid; Signal; Disease mutation;  
 CC KM Polymorphism; 3D-structure.  
 CC FT SIGNAL 1 26  
 CC FT CHAIN 27 146 CYSTATIN C.  
 CC FT ACT\_SITE 37 37 REACTIVE SITE.  
 CC FT SITE 81 85 SECONDARY AREA OF CONTACT.  
 CC FT DISULFID 99 109  
 CC FT DISULFID 123 143  
 CC  
 CC Query Match 37.9%; Score 103; DB 1; Length 146;  
 CC Best Local Similarity 40.0%; Pred. No. 3e-06;  
 CC Matches 18; Conservative 12; Mismatches 15; Indels 0; Gaps 0;  
 CC  
 CC QY 2 QYNKESDDKHYFRIFRYLKVQGVQTHLEHNLNEMQWTCQKPE 46  
 CC Db 59 EYKASNDMYHSRALQVVRARQIVAGVNFYFLVELGRTTCTRTQ 103  
 CC  
 CC RESULT 12  
 CC CYTC\_MACMU STANDARD; PRT; 146 AA.  
 CC AC 019092;  
 CC DT 15-JUL-1998 (Rel. 36, Created)  
 CC DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Cystatin C precursor.  
 CC GN CSN3.  
 CC OS Macaca mulatta (Rhesus macaque).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 CC OC Cercopithecoidea; Macaca.  
 CC OX NCBI\_TaxID=9544;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=97054523; PubMed=8898820;  
 CC RA Wei L.H., Walker L.C., Levy E.,  
 RT "Cystatin C. Icelandic-like mutation in an animal model of  
 RT cerebrovascular beta-amyloidosis.",  
 RL Stroke 27:2080-2085(1996).  
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is  
 CC thought to serve an important physiological role as a local  
 CC regulator of this enzyme activity.  
 CC -1- SIMILARITY: Belongs to the cystatin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U51912; AAB64050.1; -.  
DR HSSP; P01034; 1G96.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00003; cystatin; 1.  
DR SMART; SM00043; CY; 1.  
DR PROSITE; PS00287; CYSTATIN; 1.  
KM Thiol protease inhibitor; Amyloid; Signal.  
FT SIGNAL 1 26 BY SIMILARITY.  
FT CHAIN 27 146 CYSTATIN C.  
FT ACT SITE 37 37 REACTIVE SITE.  
FT SITE 81 85 SECONDARY AREA OF CONTACT.  
FT DISULFID 99 109 BY SIMILARITY.  
FT DISULFID 123 143 BY SIMILARITY.  
SQ SEQUENCE 146 AA; 15857 MW; F0B3B8774A29DF26 CRC64;

Query Match 37.9%; Score 103; DB 1; Length 146;  
Best Local Similarity 40.0%; Pred. No. 3e-06;  
Matches 18; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRFVYLKQVOTDHLLEYHLNEMQWTTCKPE 46  
Db 59 EYNKASNDKTHSRVRLVYRKQVVGAVNPLDVELGRITCTKIQ 103

RESULT 13  
CYTC\_BOVIN STANDARD; PRT; 148 AA.  
AC P01035;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin C precursor (Colostrum thiol proteinase inhibitor).  
GN CST3.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A., SEQUENCE OF 66-83, AND CHARACTERIZATION.  
RC TISSUE=Cerebrospinal fluid, and Choroid plexus;  
RX MEDLINE=98094199; PubMed=9434110;  
RA Olsson S.-L., Ek B., Wilm M., Broberg S., Rask L., Björck I.;  
RT "Molecular cloning and N-terminal analysis of bovine cystatin C  
RT identification of a full-length N-terminal region.";  
RL Blochim. Biophys. Acta 1343:203-210(1997).  
RN [2]  
RP SEQUENCE OF 37-148.  
RX MEDLINE=85231205; PubMed=3891407;  
RA Hirado M., Tsumasawa S., Sakiyama F., Minobe M., Fujii S.;  
RT "Complete amino acid sequence of bovine colostrum low-Mr cysteine  
RT proteinase inhibitor.";  
RL FBS Lett. 186:41-45(1985).  
CC -1- FUNCTION: This is a thiol proteinase inhibitor.  
CC -1- MASS SPECTROMETRY: MW=13420; METHOD=MALDI.  
CC -1- SIMILARITY: Belongs to the cystatin family.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; Y10811; CAA71771.1; -.  
DR HSSP; P01034; 1G96.  
DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00003; cystatin; 1.  
DR SMART; SM00043; CY; 1.  
DR PROSITE; PS00287; CYSTATIN; 1.  
KM Thiol protease inhibitor; Signal; Pyroglutamate carboxylic acid.  
FT SIGNAL 1 30 PROBABLE.  
FT CHAIN 31 148 CYSTATIN C.  
FT MOD RES 31 31 PYROLIDONE CARBOXYLIC ACID (PROBABLE).  
FT ACT SITE 40 40 REACTIVE SITE.  
FT SITE 84 88 SECONDARY AREA OF CONTACT.  
FT DISULFID 102 112 BY SIMILARITY.  
FT DISULFID 126 146 BY SIMILARITY.  
SQ SEQUENCE 148 AA; 16265 MW; E8740F837CEB9F0E CRC64;

Query Match 37.9%; Score 103; DB 1; Length 148;  
Best Local Similarity 40.0%; Pred. No. 3.1e-06;  
Matches 18; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRFVYLKQVOTDHLLEYHLNEMQWTTCKPE 46  
Db 62 EFNRSNDAYQSRVVRVVRARQVSGMVFPLDVELGRITCTKTSQ 106

RESULT 14  
CYT\_COTUJA STANDARD; PRT; 116 AA.  
AC P81061;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cystatin (Egg-white cystatin).  
OS Coturnix coturnix japonica (Japanese quail).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Coturnix.  
OX NCBI\_TaxID=93934;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Egg white; PubMed=9276465;  
RX Gerhartz B., Engh R.A., Mentele R., Eckerskorn C., Torguato R.,  
RA Wiltman J., Kolb H.J., Machleidt W., Fritz H., Auerwald E.A.;  
RT "Molecular cloning, isolation and characterization of a new member of  
RT the cystatin family and its hypothetical interaction with cathepsin  
RT B.";  
RL FBS Lett. 412:551-558(1997).  
CC -1- FUNCTION: This protein binds tightly to and inhibits papain and  
CC cathepsin B.  
CC -1- SIMILARITY: Belongs to the cystatin family.  
DR HSSP; P01036; 1CEW.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00003; cystatin; 1.  
DR SMART; SM00043; CY; 1.  
DR PROSITE; PS00287; CYSTATIN; 1.  
KM Thiol protease inhibitor; Phosphorylation.  
FT ACT SITE 9 9 REACTIVE SITE.  
FT SITE 53 57 SECONDARY AREA OF CONTACT.  
FT DISULFID 71 81  
FT DISULFID 95 115  
FT MOD RES 80 80  
SQ SEQUENCE 116 AA; 13093 MW; 48248621053A2F70 CRC64;

Query Match 36.4%; Score 99; DB 1; Length 116;  
Best Local Similarity 39.5%; Pred. No. 7.6e-06;  
Matches 17; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRFVYLKQVOTDHLLEYHLNEMQWTTCKPE 44  
Db 31 EYNKASNDKTHSRVRLVYRKQVVGAVNPLDVELGRITCTKCPK 73

RESULT 15  
CSTL\_HUMAN STANDARD; PRT; 165 AA.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 29.3611 Seconds  
(without alignments)  
526.560 Million cell updates/sec

Title: US-09-941-314-13

Perfect score: 272  
Sequence: 1 DQYKESDDKXHPRIFRVLK.....EYHLNVMQWTTQCKPPTTN 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPRMBL 25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacterioplasmid:\*  
17: sp\_archaeoplasmid:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	272	100.0	138	4	Q8WUX6
2	162	59.6	139	11	Q8K5A3
3	128	47.1	103	4	Q8WUX5
4	122	44.9	141	11	Q9DAP1
5	122	44.9	141	11	Q80ZNS
6	108	39.7	140	11	Q9EPX9
7	97	35.7	112	13	Q9BSR4
8	97	35.7	112	13	Q9BSR3
9	84	30.9	140	11	Q80Y72
10	83	30.5	425	3	Q12700
11	81	29.8	109	5	Q9TYCS
12	81	29.8	148	5	Q9NH95
13	80	29.4	161	5	Q16159
14	79	29.0	146	11	Q8K397
15	79	29.0	149	11	Q9D1B1
16	78	28.7	149	11	Q8VHC1

17	77	28.3	498	5	Q16454
18	74	27.2	144	13	Q8JFUS
19	74	27.2	148	11	Q9JMR4
20	72	26.5	125	5	Q25620
21	71	26.1	167	4	Q724J8
22	68	25.0	157	5	Q17108
23	67	24.6	127	5	P90698
24	67	24.6	284	16	Q88207
25	64.5	23.7	443	5	Q8ICX8
26	64	23.5	81	6	Q29212
27	64	23.5	167	11	Q9QWL5
28	63.5	23.3	787	11	Q8B182
29	63	23.2	438	3	Q9URJ8
30	62.5	23.0	302	4	Q9H740
31	62.5	23.0	421	4	Q9NXS0
32	62.5	23.0	995	4	Q9HC18
33	62	22.8	693	11	Q8K145
34	61.5	22.6	587	9	Q859J0
35	61.5	22.6	587	9	Q859J0
36	61	22.4	159	4	Q8TD53
37	60.5	22.2	133	5	Q8WVB6
38	60	22.1	204	4	Q8TC17
39	60	22.1	275	4	Q8WYG2
40	60	22.1	357	4	Q72465
41	60	22.1	462	13	Q72Y91
42	60	22.1	462	13	Q75YH2
43	60	22.1	465	13	Q801E5
44	59.5	21.9	128	11	Q9D1N8
45	59.5	21.9	586	5	Q94510

## ALIGNMENTS

RESULT 1	Q8WUX6	PRELIMINARY;	PRT;	138 AA.
AC	Q8WUX6	Q8WUX6		
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	SC13.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RA	Hamil K.G., Liu Q., Zhang Y.-L., French P.S., Hall S.H.;			
RT	"SC13: A novel epididymal specific member of the cystatin family."			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF35480; AAL7191.1; -			
DR	GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR000010; Cystatin.			
DR	Pfam; PF00003; Cystatin; 1.			
DR	SMART; SM00043; CY; 1.			
DR	SEQUENCE 138 AA; 16506 MW; E49440ACA3585C64 CRC64;			
QY	Query Match	100.0%;	Score 272;	DB 4; Length 138;
Db	Best Local Similarity	100.0%;	Pred. No. 7.2e-28;	
	Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 DQYKESDDKXHPRIFRVLK...YHNLNVMQWTTQCKPPTTN 49			
Db	53 DQYKESDDKXHPRIFRVLK...YHNLNVMQWTTQCKPPTTN 101			
RESULT 2	Q8K5A3	PRELIMINARY;	PRT;	139 AA.
ID	Q8K5A3			
AC	Q8K5A3			
DT	01-OCT-2002 (TREMBLrel. 22, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			

```

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Cystatin 11.
GN CS711.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Hamil K.G., Hall S.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF501290; AAM21709.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; Cy; 1.
SQ SEQUENCE 139 AA; 1686 MW; E1E36DB786B4D08C CRC64;

Query Match 59.6%; Score 162; DB 11; Length 139;
Best Local Similarity 56.5%; Pred. No. 2e-13;
Matches 26; Conservative 14; Mismatches 6; Indels 0; Gaps 0;

QY 1 DQYKESDDKXHFRIFRVLKQROVTDHLEYNLVEMQWTTCKPE 46
Db 53 EYVKKSEDDLVNFRILRIKLEKQMTNMEFHITVEMQRTTCLKTE 98

RESULT 3
Q8MXU5 PRELIMINARY; PRT; 103 AA.
AC Q8MXU5;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE SC13delta.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13, A novel epididymal specific member of the cystatin family.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF35481; AAL71992.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; Cy; 1.
SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;

Query Match 47.1%; Score 128; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYKESDDKXHFRIFRVLKQRO 24
Db 53 DQYKESDDKXHFRIFRVLKQRO 76

RESULT 4
Q9DAP1 PRELIMINARY; PRT; 141 AA.
AC Q9DAP1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE 1700006C19Rik protein.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai U., Shinesawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,
RA Atkawa T., Hara A., Fukunishi Y., Konno H., Adachi U., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirni L.W., Staudl F., Suzuki R., Tomita M., Wagner U., Maehiro T.,
RA Sakai K., Okido T., Futuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK005665; BAB24175.1; -
DR HSSP; P01038; ICEW.
DR MGD; MGI:1916544; 1700006C19Rik.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; Cy; 1.
SQ SEQUENCE 141 AA; 16811 MW; C20FA0D8B1AC378C CRC64;

Query Match 44.9%; Score 122; DB 11; Length 141;
Best Local Similarity 51.2%; Pred. No. 3.6e-08;
Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 2 QYKESDDKXHFRIFRVLKQROVTDHLEYNLVEMQWTTCK 44
Db 54 EYKASNDLVNFRVLDKQSEQITDSLEYLVNARIATWCK 96

RESULT 5
Q80ZNS PRELIMINARY; PRT; 141 AA.
AC Q80ZNS;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE RIKEN cDNA 1700006C19 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strauberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048681; AAH48681.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C/M.
DR Pfam; PF00031; Cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; Cy; 1.
DR SMART; SM00043; Cy; 1.
SQ SEQUENCE 141 AA; 16825 MW; C20FA0D8A84951F CRC64;

Query Match 44.9%; Score 122; DB 11; Length 141;
Best Local Similarity 51.2%; Pred. No. 3.6e-08;
Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 2 QYKESDDKXHFRIFRVLKQROVTDHLEYNLVEMQWTTCK 44

```

Db 54 EYKASNDLYNFRVVDLKSQEQITDLSLEYLVENIARTWCK 96

RESULT 6  
Q98PX9 PRELIMINARY; PRT; 140 AA.

AC Q98PX9; 01-JUN-2001 (TrEMBLrel. 16, Created)  
DT 01-JUN-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Cystatin C.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C;  
RX MEDLINE=21010502; PubMed=11144350;  
RA Taupin P.J., Ray J., Fischer W.H., Suh S.T., Hakanson K., Grubb A.,  
Gage F.H.;  
RT "FGF-2-Responsive neural stem cell proliferation requires CCG, a novel  
autocrine/paracrine cofactor.";  
RL Neuron 28:385-397(2000).  
DR HSSP; P01034; 1G96.  
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00031; cystatin; 1.  
DR SMART; SM00043; CY; 1.  
DR PROSITE; PS00287; CYSTATIN; 1.  
FT CHAIN 21 140 CYSTATIN C.  
FT VARIANT 16 16 A -> G.  
FT VARIANT 84 84 L -> F.  
SQ SEQUENCE 140 AA; 15517 MW; 3A563406D58D785 CRC64;  
Query Match 39.7%; Score 108; DB 11; Length 140;  
Best Local Similarity 43.5%; Pred. No. 2.4e-06;  
Matches 20; Conservative 11; Mismatches 15; Indels 0; Gaps 0;  
QY 2 QYNKESDDKXHFRIFRVLKQVROVTDHLEYHLNVEQMWTQCK 47  
Db 53 EYKASNDAYHSRAIQVVRARAKOLVAGVNYFLDVGRTTCTKSQT 98  
RESULT 7  
Q98SR4 PRELIMINARY; PRT; 112 AA.  
AC Q98SR4; 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Cystatin (Fragment).  
OS Acipenser sinensis (Chinese sturgeon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;  
OC Acipenser.  
OC NCBI\_TaxID=61970;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Bai J., Lao H., Ye X., Li Y., Lou J.;  
RT "Molecular cloning and sequence analysis of cystatin cDNA from two  
species of sturgeons.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DE EMBL; AF334610; AAK16731.1; -.  
DR HSSP; P01038; 1A90.  
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00031; cystatin; 1.  
DR SMART; SM00043; CY; 1.  
DR PROSITE; PS00287; CYSTATIN; 1.

FT NON TER 1 1  
SQ SEQUENCE 112 AA; 12231 MW; 48CEBFE8A08C00 CRC64;

Query Match 35.7%; Score 97; DB 13; Length 112;  
Best Local Similarity 41.9%; Pred. No. 5.4e-05;  
Matches 18; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKQVROVTDHLEYHLNVEQMWTQCK 44  
Db 26 EYKASNDMYIHRVSKVKVQKOVVAGIKYIVVOMGRTSCRK 68

## RESULT 8

Q98SR3 PRELIMINARY; PRT; 112 AA.

AC Q98SR3; 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Cystatin (Fragment).  
OS Acipenser schrenckii (Amar sturgeon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;  
OC Acipenser.  
OC NCBI\_TaxID=111304;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Bai J., Lao H., Ye X., Li Y., Lou J.;  
RT "Molecular cloning and sequence analysis of cystatin cDNA from two  
species of sturgeons.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DE EMBL; AF334611; AAK16732.1; -.  
DR HSSP; P01038; 1A90.  
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.  
DR InterPro; IPR000010; Cystatin.  
DR Pfam; PF00031; cystatin; 1.  
DR PROSITE; PS00287; CYSTATIN; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 112 AA; 12231 MW; 48CEBFE8A08C00 CRC64;  
Query Match 35.7%; Score 97; DB 13; Length 112;  
Best Local Similarity 41.9%; Pred. No. 5.4e-05;  
Matches 18; Conservative 13; Mismatches 12; Indels 0; Gaps 0;  
QY 2 QYNKESDDKXHFRIFRVLKQVROVTDHLEYHLNVEQMWTQCK 44  
Db 26 EYKASNDMYIHRVSKVKVQKOVVAGIKYIVVOMGRTSCRK 68  
RESULT 9  
Q80Y72 PRELIMINARY; PRT; 140 AA.  
AC Q80Y72; 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Cystatin-like 1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testicle;  
RX MEDLINE=2388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S., Wang J., Helel P.,  
Ratetsky L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Uudin T.B., Toehiyuki S., Carninci P., Prange C.,



Qy 4 NKESDDKHFRIFRVLKVGROVTDHLEHYLVNEMQWTTCK 44  
 DB 62 NQOSNDAYHMPKIKLVKSSQVAGVGMKXKKEIQVARSDCCK 102

## RESULT 13

ID 016159 PRELIMINARY; PRT; 161 AA.  
 AC 016159;  
 DT 01-JUN-1998 (TEMBLrel. 05, Created)  
 DT 01-JUN-1998 (TEMBLrel. 05, last sequence update)  
 DT 01-JUN-2003 (TEMBLrel. 24, last annotation update)  
 DE Cystatin-type cysteine proteinase inhibitor.  
 GN BM-CPI-2  
 OS Brugia malayi (Filarial nematode worm).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Brugia.  
 OX NCBI\_TaxID=6279;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gregory W.F., Blaxter M.L., Maizels R.M.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gregory W.F., Maizels R.M.;  
 RT "Two distinct cystatin-type cysteine protease inhibitors from the  
 RT parasitic nematode Brugia malayi";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AF015263; AAB69857.1; -  
 DR EMBL, AF171793; AAD51086.1; -  
 DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 DR PROSITE; PS00287; CYSTATIN; 1.  
 SQ SEQUENCE 161 AA; 18406 MW; 8081351BEE226B5 CRC64;

Query Match 29.4%; Score 80; DB 5; Length 161;  
 Best Local Similarity 34.1%; Pred. No. 0.013;  
 Matches 14; Conservative 13; Mismatches 14; Indels 0; Gaps 0;  
 Qy 4 NKESDDKHFRIFRVLKVGROVTDHLEHYLVNEMQWTTCK 44  
 DB 73 NQOSNDAYHMPKIKLVKSSQVAGVGMKXKKEIQVARSDCCK 113

RESULT 14  
 ID 08K397 PRELIMINARY; PRT; 146 AA.  
 AC 08K397;  
 DT 01-OCT-2002 (TEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TEMBLrel. 22, last sequence update)  
 DT 01-JUN-2003 (TEMBLrel. 24, last annotation update)  
 DE RIKEN cDNA 1110017B11 gene (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Straussberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, BC027680; AAH27680.1; -  
 DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 DR SMART; SM00043; CY; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 146 AA; 16380 MW; 9D77BB9A6063A5C4 CRC64;

Query Match 29.0%; Score 79; DB 11; Length 146;  
 Best Local Similarity 33.3%; Pred. No. 0.016;

Matches 14; Conservative 13; Mismatches 15; Indels 0; Gaps 0;  
 Qy 3 YNKESDDKHFRIFRVLKVGROVTDHLEHYLVNEMQWTTCK 44  
 DB 56 YNMGSDSLYYFRDVKYIDAKYQLVAGIKYVLTLDISTECRK 97

## RESULT 15

ID 09D1B1 PRELIMINARY; PRT; 149 AA.  
 AC 09D1B1;  
 DT 01-JUN-2001 (TEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, last annotation update)  
 DE 1110017E1RIK protein (Cystatin M/E) (Cystatin N homolog).  
 GN 1110017E1RIK OR CST6.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barin G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guenrich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Monbaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschki S.,  
 RA Hayaishiaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129Sv/SvTac; TISSUE=Spleen;  
 RA Zeewen P.L.J.M., van Vlijmen-Willems I.M.J.J., Hendriks W.,  
 RA Merks G.F., Schalkwijk J.;  
 RT "A mouse cystatin M/B-null mutation.";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL, AK003744; BAB22976.1; -  
 DR EMBL, AY093591; AMN1475.1; -  
 DR EMBL, AK078116; BAC37132.1; -  
 DR HSSP; P01036; 1CEW.  
 DR MGD; MGI:1920970; Cst6.  
 DR GO: GO:0001533; C:confined envelope; IDA.  
 DR GO: GO:0008544; P:epidermal differentiation; IMP.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 SQ SEQUENCE 149 AA; 16796 MW; E713EB920B0FC5 CRC64;

Query Match 29.0%; Score 79; DB 11; Length 149;  
 Best Local Similarity 33.3%; Pred. No. 0.017;

Matches 14; Conservative 13; Mismatches 15; Indels 0; Gaps 0;  
 Qy 3 YNKSDDKHYFRIFRVKQVOTDHLDEYHUNVEMQWTCOK 44  
 Db 59 YNMGSDSLYYFRDIXVIDAKYQLVAGIKYITLDIESTECKR 100

Search completed: March 18, 2004, 14:21:31  
 Job time : 29.3611 secs

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

